

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 07:06:01 ; Search time 22.25 Seconds
(without alignments)
59.612 Million cell updates/sec

Title: US-10-066-965A-2

Perfect score: 20

Sequence: 1 QWSSWALGWRLLRRYGWGM 20

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 478139 seqs, 66318000 residues

Word size : 0

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PTUS COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	65.0	20	3	US-08-504-538A-12
2	13	65.0	20	3	US-08-630-052-12
3	13	65.0	20	5	PCT-US95-09307-12
4	8	40.0	480	4	US-09-252-991A-31470
5	6	30.0	135	4	US-09-270-767-31648
6	6	30.0	135	4	US-09-270-767-46865
7	6	30.0	138	4	US-09-252-991A-32834
8	6	30.0	154	3	US-09-193-104-26
9	6	30.0	171	4	US-09-252-991A-26818
10	6	30.0	277	4	US-09-247-890-16
11	6	30.0	282	4	US-09-247-890-16
12	6	30.0	282	4	US-09-724-852-16
13	6	30.0	282	4	US-09-724-852-16
14	6	30.0	346	5	PCT-US96-10602-2
15	6	30.0	395	4	US-09-489-039A-7940
16	6	30.0	421	4	US-09-248-796A-20409
17	6	30.0	478	4	US-09-252-991A-22078
18	6	30.0	800	4	US-09-252-991A-24635
19	6	30.0	939	4	US-09-854-845-16
20	6	30.0	954	4	US-09-854-845-14
21	6	30.0	1034	4	US-09-854-845-6
22	6	30.0	1049	4	US-09-854-845-2
23	6	30.0	1078	4	US-09-854-845-8
24	6	30.0	1093	4	US-09-854-845-4
25	6	30.0	1126	4	US-09-854-845-12
26	6	30.0	1151	4	US-09-854-845-10
27	6	30.0	1509	4	US-09-676-519-27

28	5	25.0	12	4	US-08-988-024C-7	Sequence 7, Appli
29	5	25.0	19	1	US-07-944-143C-21	Sequence 21, Appl
30	5	25.0	19	5	PCT-US93-08214-21	Sequence 21, Appl
31	5	25.0	20	4	US-09-155-613A-6	Sequence 6, Appli
32	5	25.0	20	4	US-09-498-134A-3	Sequence 3, Appli
33	5	25.0	21	1	US-07-944-143C-20	Sequence 20, Appl
34	5	25.0	21	5	PCT-US93-08214-20	Sequence 20, Appl
35	5	25.0	24	4	US-09-155-613A-97	Sequence 97, Appl
36	5	25.0	25	4	US-09-155-613A-1	Sequence 1, Appli
37	5	25.0	25	4	US-09-155-613A-98	Sequence 98, Appl
38	5	25.0	35	4	US-09-155-613A-22	Sequence 22, Appl
39	5	25.0	35	4	US-09-155-613A-23	Sequence 23, Appl
40	5	25.0	38	1	US-08-179-632-7	Sequence 7, Appli
41	5	25.0	38	1	US-08-440-174A-7	Sequence 7, Appli
42	5	25.0	38	5	PCT-US95-00062-7	Sequence 7, Appli
43	5	25.0	51	4	US-09-621-976-4432	Sequence 432, Ap
44	5	25.0	60	5	PCT-US94-14074-1	Sequence 1, Appli
45	5	25.0	61	3	US-09-134-001C-3836	Sequence 3836, Ap
46	5	25.0	62	4	US-09-513-999C-4436	Sequence 4436, Ap
47	5	25.0	73	4	US-09-252-991A-29973	Sequence 29973, A
48	5	25.0	76	4	US-08-513-999C-6143	Sequence 6143, Ap
49	5	25.0	79	4	US-09-198-452A-1187	Sequence 1187, Ap
50	5	25.0	79	4	US-09-270-767-59078	Sequence 59078, A
51	5	25.0	82	4	US-09-205-258-403	Sequence 403, App
52	5	25.0	88	4	US-09-621-976-3972	Sequence 3972, Ap
53	5	25.0	92	4	US-09-673-809-25	Sequence 25, Appl
54	5	25.0	98	4	US-09-652-345-4	Sequence 4, Appli
55	5	25.0	98	4	US-09-540-236-2144	Sequence 2144, Ap
56	5	25.0	108	4	US-09-710-279-1724	Sequence 1724, Ap
57	5	25.0	113	4	US-09-621-976-7483	Sequence 7483, Ap
58	5	25.0	119	4	US-09-621-976-5186	Sequence 5186, Ap
59	5	25.0	119	4	US-09-621-976-5890	Sequence 5890, Ap
60	5	25.0	119	4	US-09-513-999C-5792	Sequence 5792, Ap
61	5	25.0	123	4	US-09-816-248-2	Sequence 2, Appli
62	5	25.0	123	4	US-09-816-248-3	Sequence 3, Appli
63	5	25.0	123	4	US-09-816-248-4	Sequence 4, Appli
64	5	25.0	129	3	US-08-980-523-11	Sequence 11, Appl
65	5	25.0	131	4	US-09-540-236-3251	Sequence 3251, Ap
66	5	25.0	137	3	US-09-036-574-4	Sequence 4, Appli
67	5	25.0	137	4	US-08-454-294A-4	Sequence 4, Appli
68	5	25.0	144	4	US-09-270-767-32020	Sequence 32020, A
69	5	25.0	144	4	US-09-270-767-47237	Sequence 47237, A
70	5	25.0	158	4	US-09-252-991A-27239	Sequence 27239, A
71	5	25.0	168	4	US-09-252-991A-20364	Sequence 20364, A
72	5	25.0	168	4	US-09-252-991A-30988	Sequence 30988, A
73	5	25.0	169	4	US-09-252-991A-32019	Sequence 32019, A
74	5	25.0	170	4	US-09-252-991A-18579	Sequence 18579, A
75	5	25.0	173	4	US-09-252-991A-20339	Sequence 20339, A
76	5	25.0	176	4	US-09-252-991A-24281	Sequence 24281, A
77	5	25.0	180	4	US-09-248-796A-17721	Sequence 17721, A
78	5	25.0	182	1	US-08-127-954-135	Sequence 135, App
79	5	25.0	182	1	US-08-127-954-137	Sequence 137, App
80	5	25.0	182	1	US-08-127-954-139	Sequence 139, App
81	5	25.0	182	1	US-08-127-954-140	Sequence 140, App
82	5	25.0	182	1	US-08-127-954-141	Sequence 141, App
83	5	25.0	182	1	US-08-127-954-142	Sequence 142, App
84	5	25.0	182	1	US-08-127-954-143	Sequence 143, App
85	5	25.0	182	1	US-08-127-954-144	Sequence 144, App
86	5	25.0	182	1	US-08-127-954-145	Sequence 145, App
87	5	25.0	182	1	US-08-127-954-146	Sequence 146, App
88	5	25.0	182	1	US-08-127-954-147	Sequence 147, App
89	5	25.0	182	1	US-08-127-954-148	Sequence 148, App
90	5	25.0	182	1	US-08-127-954-149	Sequence 149, App
91	5	25.0	182	1	US-08-127-954-153	Sequence 153, App
92	5	25.0	182	1	US-08-127-954-154	Sequence 154, App
93	5	25.0	182	1	US-08-127-954-155	Sequence 155, App
94	5	25.0	182	1	US-08-127-954-156	Sequence 156, App
95	5	25.0	182	1	US-08-127-954-157	Sequence 157, App
96	5	25.0	182	1	US-08-127-954-158	Sequence 158, App
97	5	25.0	182	1	US-08-127-954-159	Sequence 159, App
98	5	25.0	182	1	US-08-127-954-160	Sequence 160, App
99	5	25.0	182	1	US-08-127-954-161	Sequence 161, App
100	5	25.0	182	1		

ALIGNMENTS

```

RESULT 1
US-08-504-538A-12
; Sequence 12, Application US/08504538A
; Patent No. 6004746
; GENERAL INFORMATION:
; APPLICANT: Brent, Roger
; APPLICANT: McCoy, John M.
; APPLICANT: Jessen, Timm H.
; TITLE OF INVENTION: INTERACTION TRAP SYSTEMS FOR DETECTING
; TITLE OF INVENTION: PROTEIN INTERACTIONS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2214
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,538A
; FILING DATE: 07/20/95
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/278,082
; FILING DATE: 07/20/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/259001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 428-0200
; TELEFAX: (617) 428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
US-08-504-538A-12

Query Match 65.0%; Score 13; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 WALGWRWLRRYGW 18
Db 6 WALGWRWLRRYGW 18

RESULT 2
US-08-630-052-12
; Sequence 12, Application US/08630052
; Patent No. 639296
; GENERAL INFORMATION:
; APPLICANT: Brent, Roger
; APPLICANT: McCoy, John M.
; APPLICANT: Jessen, Timm H.
; APPLICANT: Xu, Chanxing Wilson
; TITLE OF INVENTION: INTERACTION TRAP SYSTEMS FOR DETECTING PROTEIN
; TITLE OF INVENTION: INTERACTIONS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,052
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/504,538
; FILING DATE: July 20, 1995
; APPLICATION NUMBER: 08/278,082
; FILING DATE: July 20, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Karen F. Lech
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/311001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
US-08-630-052-12

Query Match 65.0%; Score 13; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 WALGWRWLRRYGW 18
Db 6 WALGWRWLRRYGW 18

RESULT 3
PCT-US95-09307-12
; Sequence 12, Application PC/TUS9509307
; GENERAL INFORMATION:
; APPLICANT: Brent, Roger
; APPLICANT: McCoy, John M.
; APPLICANT: Jessen, Timm H.
; APPLICANT: Xu, Chanxing Wilson
; TITLE OF INVENTION: INTERACTION TRAP SYSTEMS FOR
; TITLE OF INVENTION: DETECTING PROTEIN INTERACTIONS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version
; SOFTWARE: #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09307
; FILING DATE:
; CLASSIFICATION:

```



```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/288001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US95-09307-12

Query Match 55.0%; Score 13; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 WALGWRWLRYYGW 18
Db 6 WALGWRWLRYYGW 18

RESULT 4
US-09-252-991A-31470
; Sequence 31470, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31470
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31470

Query Match 40.0%; Score 8; DB 4; Length 480;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LGWRWLR 15
Db 255 LGWRWLR 262

RESULT 5
US-09-270-767-31648
; Sequence 31648, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31648
; LENGTH: 135
```

```
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-31648

Query Match 30.0%; Score 6; DB 4; Length 135;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLR 15
Db 24 WRWLR 29

RESULT 6
US-09-270-767-46865
; Sequence 46865, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46865
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46865

Query Match 30.0%; Score 6; DB 4; Length 135;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLR 15
Db 24 WRWLR 29

RESULT 7
US-09-252-991A-32834
; Sequence 32834, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32834
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32834

Query Match 30.0%; Score 6; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLR 15
Db 49 WRWLR 54

RESULT 8
US-09-193-104-26
```

```
; Sequence 26, Application US/09193104A
; Patent No. 6172193
; GENERAL INFORMATION:
; APPLICANT: Primi, Daniele
; APPLICANT: Fiordalisi, Gianfranco
; APPLICANT: Palla, Mario
; TITLE OF INVENTION: Escape Mutant of the Surface Antigen of Hepatitis B
; TITLE OF INVENTION: Virus
; FILE REFERENCE: SBD1004US
; CURRENT APPLICATION NUMBER: US/09/193,104A
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: EP 97830635.5
; EARLIER FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-09-193-104-26

Query Match          30.0%; Score 6; DB 3; Length 154;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 SSWALG 9
Db      121 SSWALG 126

RESULT 9
US-09-252-991A-26818
; Sequence 26818, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26818
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26818

Query Match          30.0%; Score 6; DB 4; Length 171;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      10 WRWLR 15
Db      8 WRWLR 13

RESULT 10
US-09-252-991A-17567
; Sequence 17567, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1999-02-18
```

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17567
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17567

Query Match          30.0%; Score 6; DB 4; Length 277;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      10 WRWLR 15
Db      265 WRWLR 270

RESULT 11
US-09-247-890-16
; Sequence 16, Application US/09247890
; Patent No. 6541011
; GENERAL INFORMATION:
; APPLICANT: Punnonen, Juha
; APPLICANT: Bass, Steven H.
; APPLICANT: Whalen, Robert Gerald
; APPLICANT: Howard, Russell
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Maxygen, Inc.
; TITLE OF INVENTION: Antigen Library Immunization
; FILE REFERENCE: 018097-028710US
; CURRENT APPLICATION NUMBER: US/09/247,890
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: US 60/074,294
; EARLIER FILING DATE: 1998-02-11
; EARLIER APPLICATION NUMBER: US 60/105,509
; EARLIER FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Woodchuck hepatitis B virus
US-09-247-890-16

Query Match          30.0%; Score 6; DB 4; Length 282;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 SSWALG 9
Db      210 SSWALG 215

RESULT 12
US-09-724-969-16
; Sequence 16, Application US/09724969
; Patent No. 6569435
; GENERAL INFORMATION:
; APPLICANT: Punnonen, Juha
; APPLICANT: Bass, Steven H.
; APPLICANT: Whalen, Robert Gerald
; APPLICANT: Howard, Russell
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Maxygen, Inc.
; TITLE OF INVENTION: Antigen Library Immunization
; FILE REFERENCE: 018097-028710US
; CURRENT APPLICATION NUMBER: US/09/724,969
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/247,890
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: US 60/105,509
```

; PRIOR FILING DATE: 1998-10-23
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 16
 ; LENGTH: 282
 ; TYPE: PRT
 ; ORGANISM: Woodchuck hepatitis B virus
 US-09-724-969-16

Query Match 30.0%; Score 6; DB 4; Length 282;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SSWALG 9
 Db 210 SSWALG 215
 |||||

RESULT 13
 US-09-724-852-16
 ; Sequence 16, Application US/09724852
 ; Patent No. 6576757
 ; GENERAL INFORMATION:
 ; APPLICANT: Punnonen, Juha
 ; APPLICANT: Bass, Steven H.
 ; APPLICANT: Whalen, Robert Gerald
 ; APPLICANT: Howard, Russell
 ; APPLICANT: Stemmer, Willem P.C.
 ; APPLICANT: Maxygen, Inc.
 ; TITLE OF INVENTION: Antigen Library Immunization
 ; FILE REFERENCE: 018097-028710US
 ; CURRENT APPLICATION NUMBER: US/09724,852
 ; CURRENT FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: US/09/247,890
 ; PRIOR FILING DATE: 1999-02-10
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/074,294
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-11
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/105,509
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-23
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 16
 ; LENGTH: 282
 ; TYPE: PRT
 ; ORGANISM: Woodchuck hepatitis B virus
 US-09-724-852-16

Query Match 30.0%; Score 6; DB 4; Length 282;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SSWALG 9
 Db 210 SSWALG 215
 |||||

RESULT 14
 PCT-US96-10602-2
 ; Sequence 2, Application PC/TUS9610602
 ; GENERAL INFORMATION:
 ; APPLICANT: The General Hospital Corporation
 ; TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US96/10602
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA: 60/017,814
 ; FILING DATE: 20-JUN-1995
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Clark, Paul T.
 ; REGISTRATION NUMBER: 30,162
 ; REFERENCE/DOCKET NUMBER: 00786/282001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 346 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 PCT-US96-10602-2

Query Match 30.0%; Score 6; DB 5; Length 346;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SSWALG 9
 Db 274 SSWALG 279
 |||||

RESULT 15
 US-09-489-039A-7940
 ; Sequence 7940, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 7940
 ; LENGTH: 395
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-7940

Query Match 30.0%; Score 6; DB 4; Length 395;
 Best Local Similarity 100.0%; Pred. No. 51;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRYYG 17
 Db 388 WLRYYG 393
 |||||

RESULT 16
 US-09-248-796A-20409
 ; Sequence 20409, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
 ; FILE REFERENCE: 107196.132

;; CURRENT APPLICATION NUMBER: US/09/248,796A
;; CURRENT FILING DATE: 1999-02-12
;; PRIOR APPLICATION NUMBER: US 60/074,725
;; PRIOR FILING DATE: 1998-02-13
;; PRIOR APPLICATION NUMBER: US 60/096,409
;; PRIOR FILING DATE: 1998-08-13
;; NUMBER OF SEQ ID NOS: 28208
;; SEQ ID NO 20409
;; LENGTH: 421
;; TYPE: PRT
;; ORGANISM: Candida albicans
US-09-248-796A-20409

Query Match 30.0%; Score 6; DB 4; Length 421;
Best Local Similarity 100.0%; Pred.No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ALGWRW 12
Db 129 ALGWRW 134
|||||

RESULT 17
US-09-252-991A-22078
; Sequence 22078, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22078
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22078

Query Match 30.0%; Score 6; DB 4; Length 478;
Best Local Similarity 100.0%; Pred.No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLR 15
Db 443 WRWLR 448
|||||

RESULT 18
US-09-252-991A-24635
; Sequence 24635, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24635
; LENGTH: 800
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-24635

Query Match 30.0%; Score 6; DB 4; Length 800;
Best Local Similarity 100.0%; Pred.No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 WRWLR 15
Db 487 WRWLR 492
|||||

RESULT 19
US-09-854-845-16
; Sequence 16, Application US/09854845
; Patent No. 6750054
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wang, Xiaoming
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6750054el Human Semaphorin Homologs and Polynucleotides Encod
; FILE REFERENCE: LEX-0177-USA
; CURRENT APPLICATION NUMBER: US/09/854,845
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/205,274
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/208,893
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 939
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-854-845-16

Query Match 30.0%; Score 6; DB 4; Length 939;
Best Local Similarity 100.0%; Pred.No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSSWAL 8
Db 473 WSSWAL 478
|||||

RESULT 20
US-09-854-845-14
; Sequence 14, Application US/09854845
; Patent No. 6750054
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wang, Xiaoming
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6750054el Human Semaphorin Homologs and Polynucleotides Encod
; FILE REFERENCE: LEX-0177-USA
; CURRENT APPLICATION NUMBER: US/09/854,845
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/205,274
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/208,893
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 954
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-854-845-14

Query Match 30.0%; Score 6; DB 4; Length 954;
Best Local Similarity 100.0%; Pred.No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSSWAL 8
|||||
Db 473 WSSWAL 478

RESULT 21

US-09-854-845-6
; Sequence 6, Application US/09854845
; Patent No. 6750054

; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade

; APPLICANT: Wang, Xiaoming

; APPLICANT: Scoville, John

; APPLICANT: Turner, C. Alexander Jr.

; TITLE OF INVENTION: No. 6750054el Human Semaphorin Homologs and Polynucleotides Encod

; FILE REFERENCE: LEX-0177-USA

; CURRENT APPLICATION NUMBER: US/09/854,845

; PRIOR FILING DATE: 2001-05-14

; PRIOR APPLICATION NUMBER: US 60/205,274

; PRIOR FILING DATE: 2000-05-18

; PRIOR APPLICATION NUMBER: US 60/208,893

; PRIOR FILING DATE: 2000-06-02

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 1034

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-854-845-6

Query Match 30.0%; Score 6; DB 4; Length 1034;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSSWAL 8
|||||
Db 568 WSSWAL 573

RESULT 22

US-09-854-845-2

; Sequence 2, Application US/09854845

; Patent No. 6750054

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Wang, Xiaoming

; APPLICANT: Scoville, John

; APPLICANT: Turner, C. Alexander Jr.

; TITLE OF INVENTION: No. 6750054el Human Semaphorin Homologs and Polynucleotides Encod

; FILE REFERENCE: LEX-0177-USA

; CURRENT APPLICATION NUMBER: US/09/854,845

; PRIOR FILING DATE: 2001-05-14

; PRIOR APPLICATION NUMBER: US 60/205,274

; PRIOR FILING DATE: 2000-05-18

; PRIOR APPLICATION NUMBER: US 60/208,893

; PRIOR FILING DATE: 2000-06-02

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 1049

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-854-845-2

Query Match 30.0%; Score 6; DB 4; Length 1049;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSSWAL 8
|||||
Db 568 WSSWAL 573

RESULT 23

US-09-854-845-8

; Sequence 8, Application US/09854845

; Patent No. 6750054

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Wang, Xiaoming

; APPLICANT: Scoville, John

; APPLICANT: Turner, C. Alexander Jr.

; TITLE OF INVENTION: No. 6750054el Human Semaphorin Homologs and Polynucleotides Encod

; FILE REFERENCE: LEX-0177-USA

; CURRENT APPLICATION NUMBER: US/09/854,845

; CURRENT FILING DATE: 2001-05-14

; PRIOR APPLICATION NUMBER: US 60/205,274

; PRIOR FILING DATE: 2000-05-18

; PRIOR APPLICATION NUMBER: US 60/208,893

; PRIOR FILING DATE: 2000-06-02

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 1078

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-854-845-8

Query Match 30.0%; Score 6; DB 4; Length 1078;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSSWAL 8
|||||
Db 612 WSSWAL 617

RESULT 24

US-09-854-845-4

; Sequence 4, Application US/09854845

; Patent No. 6750054

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Wang, Xiaoming

; APPLICANT: Scoville, John

; APPLICANT: Turner, C. Alexander Jr.

; TITLE OF INVENTION: No. 6750054el Human Semaphorin Homologs and Polynucleotides Encod

; FILE REFERENCE: LEX-0177-USA

; CURRENT APPLICATION NUMBER: US/09/854,845

; CURRENT FILING DATE: 2001-05-14

; PRIOR APPLICATION NUMBER: US 60/205,274

; PRIOR FILING DATE: 2000-05-18

; PRIOR APPLICATION NUMBER: US 60/208,893

; PRIOR FILING DATE: 2000-06-02

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 1093

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-854-845-4

Query Match 30.0%; Score 6; DB 4; Length 1093;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSSWAL 8
|||||
Db 612 WSSWAL 617

RESULT 25

US-09-854-845-12

; Sequence 12, Application US/09854845

; Patent No. 6750054

```
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wang, Xiaoming
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 675005461 Human Semaphorin Homologs and Polynucleotides Encod
; FILE REFERENCE: LEX-0177-USA
; CURRENT APPLICATION NUMBER: US/09/854,845
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/205,274
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/208,893
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1136
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-854-845-12

Query Match          30.0%; Score 6; DB 4; Length 1136;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 WSSWAL 8
DB      670 WSSWAL 675

RESULT 26
US-09-854-845-10
; Sequence 10, Application US/09854845
; Patent No. 6750054
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wang, Xiaoming
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 675005461 Human Semaphorin Homologs and Polynucleotides Encod
; FILE REFERENCE: LEX-0177-USA
; CURRENT APPLICATION NUMBER: US/09/854,845
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/205,274
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/208,893
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-854-845-10

Query Match          30.0%; Score 6; DB 4; Length 1151;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 WSSWAL 8
DB      670 WSSWAL 675

RESULT 27
US-09-676-519-27
; Sequence 27, Application US/09676519
; Patent No. 6737508
; GENERAL INFORMATION:
; APPLICANT: PELLETIER, JERRY
; APPLICANT: GROS, PHILIPPE
; APPLICANT: DUBOW, MICHAEL
; TITLE OF INVENTION: DNA SEQUENCES FROM STAPHYLOCOCCUS AUREUS BATERIOPHAGES
```

```
; TITLE OF INVENTION: 3A, 77, AND 96 THAT ENCODE ANTI-MICROBIAL POLYPEPTIDES
; FILE REFERENCE: 073406-0404
; CURRENT APPLICATION NUMBER: US/09/676,519
; CURRENT FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/407,804
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: 60/110,992
; PRIOR FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 1509
; TYPE: PRT
; ORGANISM: Staphylococcus bacteriophage
US-09-676-519-27
```

```
Query Match          30.0%; Score 6; DB 4; Length 1509;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      14 RRYGNG 19
DB      1357 RRYGNG 1362
```

```
RESULT 28
US-08-988-024C-7
; Sequence 7, Application US/08988024C
; Patent No. 6635452
; GENERAL INFORMATION:
; APPLICANT: Monforte, Joseph A.
; APPLICANT: Becker, Christopher H.
; APPLICANT: Pollart, Daniel J.
; APPLICANT: Shaler, Thomas A.
; TITLE OF INVENTION: Releasable No. 6635452volatile Mass-Label Molecules
; FILE REFERENCE: 24736-2057
; CURRENT APPLICATION NUMBER: US/08/988,024C
; CURRENT FILING DATE: 1997-12-10
; PRIOR APPLICATION NUMBER: US 60/033,037
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: US 60/046,719
; PRIOR FILING DATE: 1997-05-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligopeptide
US-08-988-024C-7
```

```
Query Match          25.0%; Score 5; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      12 WLRRY 16
DB      5 WLRRY 9
```

```
RESULT 29
US-07-944-143C-21
; Sequence 21, Application US/07944143C
; Patent No. 5719064
; GENERAL INFORMATION:
; APPLICANT: Scofield, R. Hal
; APPLICANT: Harley, John B.
; TITLE OF INVENTION: Peptide Diagnostics and Therapeutics for
; TITLE OF INVENTION: Spontylarthropathies
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
```

```
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/944,143C
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF138
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)873-8794
; TELEFAX: (404)873-8795
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-07-944-143C-21

Query Match 25.0%; Score 5; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRY 16
Db 2 WLRRY 6

RESULT 30
PCT-US93-08214-21
; Sequence 21, Application PC/TUS9308214
; GENERAL INFORMATION:
; APPLICANT: Oklahoma Medical Research Foundation
; TITLE OF INVENTION: Peptide Diagnostics and Therapeutics for spondyloarthropat
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: United States
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08214
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF138
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; PCT-US93-08214-21

Query Match 25.0%; Score 5; DB 5; Length 19;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRY 16
Db 2 WLRRY 6

RESULT 31
US-09-155-613A-6
; Sequence 6, Application US/09155613A
; Patent No. 6420120
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/09/155,613A
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide MH20
; US-09-155-613A-6

Query Match 25.0%; Score 5; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRY 16
Db 10 WLRRY 14

RESULT 32
US-09-498-134A-3
; Sequence 3, Application US/09498134A
; Patent No. 6649396
; GENERAL INFORMATION:
; APPLICANT: Curiel, David T.
; APPLICANT: Douglas, Joanne T.
; APPLICANT: Kraenkyh, Victor N.
; APPLICANT: Dmitriev, Igor
; TITLE OF INVENTION: Fiber Receptor-Independent System for the
; TITLE OF INVENTION: Propagation of Adenoviral Vectors
```

FILE REFERENCE: D6159
CURRENT APPLICATION NUMBER: US/09/498,134A
CURRENT FILING DATE: 2000-02-03
PRIOR APPLICATION NUMBER: US 60/118,860
PRIOR FILING DATE: 1999-02-05
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 3
LENGTH: 20
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: amino acid sequence of the icosapeptide MH20;
OTHER INFORMATION: the extracellular virus-binding domain of the
OTHER INFORMATION: second artificial receptor
US-09-498-134A-3

Query Match 25.0%; Score 5; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRY 16
|||||
DB 10 WLRRY 14

RESULT 33
US-07-944-143C-20
Sequence 20, Application US/07944143C
Patent No. 5719064
GENERAL INFORMATION:
APPLICANT: Scofield, R. Hal
APPLICANT: Harley, John B.
TITLE OF INVENTION: Peptide Diagnostics and Therapeutics for
TITLE OF INVENTION: Spondyloarthropathies
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/944,143C
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMR138
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)873-8794
TELEFAX: (404)873-8795
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-07-944-143C-20

Query Match 25.0%; Score 5; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 WLRRY 16
|||||
DB 2 WLRRY 6

RESULT 34
PCT-US93-08214-20
Sequence 20, Application PC/TUS9308214
GENERAL INFORMATION:
APPLICANT: Oklahoma Medical Research Foundation
TITLE OF INVENTION: Peptide Diagnostics and Therapeutics
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: United States
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08214
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMR138
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 815-6508
TELEFAX: (404) 815-6555
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
PCT-US93-08214-20

Query Match 25.0%; Score 5; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 WLRRY 16
|||||
DB 2 WLRRY 6

RESULT 35
US-09-155-613A-97
Sequence 97, Application US/09155613A
Patent No. 6420120
GENERAL INFORMATION:
APPLICANT: Boulanger, Pierre
APPLICANT: Hong, Saw See
APPLICANT: Karayan, Lucie
TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
FILE REFERENCE: 032751-036
CURRENT APPLICATION NUMBER: US/09/155,613A


```
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Phagotopoe
US-09-155-613A-97
```

```
Query Match 25.0%; Score 5; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 12 WLRRY 16
Db 11 WLRRY 15
```

```
RESULT 36
US-09-155-613A-1
; Sequence 1, Application US/09155613A
; Patent No. 6420120
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/09/155,613A
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-155-613A-1
```

```
Query Match 25.0%; Score 5; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 12 WLRRY 16
Db 12 WLRRY 16
```

```
RESULT 37
US-09-155-613A-98
; Sequence 98, Application US/09155613A
; Patent No. 6420120
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/09/155,613A
; CURRENT FILING DATE: 1998-09-30
```

```
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Phagotopoe
US-09-155-613A-98
```

```
Query Match 25.0%; Score 5; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 12 WLRRY 16
Db 12 WLRRY 16
```

```
RESULT 38
US-09-155-613A-22
; Sequence 22, Application US/09155613A
; Patent No. 6420120
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/09/155,613A
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-155-613A-22
```

```
Query Match 25.0%; Score 5; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 12 WLRRY 16
Db 10 WLRRY 14
```

```
RESULT 39
US-09-155-613A-23
; Sequence 23, Application US/09155613A
; Patent No. 6420120
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/09/155,613A
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
```

;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: FR 97/01005
;; PRIOR FILING DATE: 1997-01-30
;; PRIOR APPLICATION NUMBER: FR 97/11166
;; PRIOR FILING DATE: 1997-09-09
;; NUMBER OF SEQ ID NOS: 98
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 23
;; LENGTH: 35
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-155-613A-23

Query Match 25.0%; Score 5; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRY 16
|||||
DB 25 WLRRY 29

RESULT 40
US-08-179-632-7
; Sequence 7, Application US/08179632
; Patent No. 5607914
; GENERAL INFORMATION:
; APPLICANT: Rao, A. Gururaj; Zhong, Lingxiu
; TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pioneer Hi-Bred International, Inc.
; STREET: 700 Capital Square, 400 Locust Street
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS/Microsoft Windows
; SOFTWARE: Microsoft Windows No. 5607914epad
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/179,632
; FILING DATE: 07-JAN-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/079,512
; FILING DATE: 06/18/93
; ATTORNEY/AGENT INFORMATION:
; NAME: Roth, Michael J.
; REGISTRATION NUMBER: 29,342
; REFERENCE/DOCKET NUMBER: 0233 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 245-3594
; TELEFAX: (515) 245-3634
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: hordothionin derivative
US-08-179-632-7

Query Match 25.0%; Score 5; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RWLRR 15
|||||
DB 8 RWLRR 12

RESULT 41
US-08-440-174A-7
; Sequence 7, Application US/08440174A
; Patent No. 5717061
; GENERAL INFORMATION:
; APPLICANT: Rao, Gururaj A.
; APPLICANT: Zhong, Lingxiu
; TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: 7100 N.W. 62nd Avenue
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,174A
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/079,512
; FILING DATE: 18-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bobrowicz, Donna
; REGISTRATION NUMBER: 32,196
; REFERENCE/DOCKET NUMBER: 0234R2D-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4896
; TELEFAX: (515) 334-6883
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-440-174A-7

Query Match 25.0%; Score 5; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RWLRR 15
|||||
DB 8 RWLRR 12

RESULT 42
PCT-US95-00062-7
; Sequence 7, Application PC/TUS9500062
; GENERAL INFORMATION:
; APPLICANT: Pioneer Hi-Bred International, Inc.
; TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pioneer Hi-Bred International, Inc.
; STREET: 700 Capital Square, 400 Locust Street
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS/Microsoft Windows

;; SOFTWARE: Microsoft Windows Notepad
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/00062
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Yates, Michael E.; Sweeney, Patricia A.;
;; NAME: Roth, Michael J.; & Simon, Soma G.
;; REGISTRATION NUMBER:
;; REFERENCE/DOCKET NUMBER: 2342-PCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (515) 248-4800
;; TELEFAX: (515) 248-4844
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; DESCRIPTION: hordothionin derivative
PCT-US95-00062-7

Query Match 25.0%; Score 5; DB 5; Length 38;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RWLR 15
Db 8 RWLR 12

RESULT 43
US-09-621-976-4432
;; Sequence 4432, Application US/09621976
;; Patent No. 6639063
;; GENERAL INFORMATION:
;; APPLICANT: Dumas Milne Edwards, J.B.
;; APPLICANT: Tobert, S.
;; APPLICANT: Giordano, J.Y.
;; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;; FILE REFERENCE: GENSET.054PR2
;; CURRENT APPLICATION NUMBER: US/09/621,976
;; CURRENT FILING DATE: 2000-07-21
;; NUMBER OF SEQ ID NOS: 19335
;; SOFTWARE: Patent.pm
;; SEQ ID NO 4432
;; LENGTH: 51
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: UNSURE
;; LOCATION: 23
;; OTHER INFORMATION: Xaa = Ala,Glu,Gly,Val
;; NAME/KEY: UNSURE
;; LOCATION: 19
;; OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr
;; NAME/KEY: UNSURE
;; LOCATION: 31
;; OTHER INFORMATION: Xaa = Ile,Asn,Ser,Thr
US-09-621-976-4432

Query Match 25.0%; Score 5; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 SWALG 9
Db 45 SWALG 49

RESULT 44
PCT-US94-14074-1
;; Sequence 1, Application PC/TUS9414074
;; GENERAL INFORMATION:
;; APPLICANT: Maksymowich, Andrew B
;; APPLICANT: Hsu, Tin-Chen
;; APPLICANT: Jameson, Bradford A
;; APPLICANT: Litwack, Gerald
;; TITLE OF INVENTION: Biologically Active Compounds
;; NUMBER OF SEQUENCES: 1
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
;; STREET: One Liberty Place 46th Floor
;; CITY: Philadelphia
;; STATE: Pennsylvania
;; COUNTRY: United States of America
;; ZIP: 19103
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US94/14074
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/164,102
;; FILING DATE: 07-DEC-1993
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Deluca, Mark
;; REGISTRATION NUMBER: 33,229
;; REFERENCE/DOCKET NUMBER: TJU-1418
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 215-568-3100
;; TELEFAX: 215-568-3439
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 60 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
PCT-US94-14074-1

Query Match 25.0%; Score 5; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ALGWR 11
Db 55 ALGWR 59

RESULT 45
US-09-134-001C-3836
;; Sequence 3836, Application US/09134001C
;; Patent No. 6380370
;; GENERAL INFORMATION:
;; APPLICANT: Lynn Doucette-Stamm et al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: GTC-007
;; CURRENT APPLICATION NUMBER: US/09/134,001C
;; CURRENT FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/064,964
;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 3836
;; LENGTH: 61

```
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3836

Query Match      25.0%; Score 5; DB 3; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      11 RWLRR 15
      |||||
Db      38 RWLRR 42

RESULT 46
US-09-513-999C-4436
; Sequence 4436, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1993-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 4436
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SIGNAL
; LOCATION: -25...-1
; OTHER INFORMATION: score 4
; OTHER INFORMATION: seq LLQELGLCMCLLS/AE
US-09-513-999C-4436

Query Match      25.0%; Score 5; DB 4; Length 62;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      10 WRWLR 14
      |||||
Db      33 WRWLR 37

RESULT 47
US-09-252-991A-29973
; Sequence 29973, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29973
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29973

Query Match      25.0%; Score 5; DB 4; Length 73;
```

```
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      11 RWLRR 15
      |||||
Db      10 RWLRR 14

RESULT 48
US-09-513-999C-6143
; Sequence 6143, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6143
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-6143

Query Match      25.0%; Score 5; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      11 RWLRR 15
      |||||
Db      16 RWLRR 20

RESULT 49
US-09-198-452A-1187
; Sequence 1187, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1187
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1187

Query Match      25.0%; Score 5; DB 4; Length 79;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 SSWAL 8
      |||||
Db      36 SSWAL 40

RESULT 50
US-09-270-767-59078
; Sequence 59078, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
```

```
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1998-03-17
; NUMBER OF SEQ ID NOS: 6217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59078
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-59078

Query Match          25.0%; Score 5; DB 4; Length 79;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LRRYG 17
Db 38 LRRYG 42

RESULT 51
US-09-205-258-403
; Sequence 403, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1998-03-17
; NUMBER OF SEQ ID NOS: 6217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59078
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (15)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-205-258-403

Query Match          25.0%; Score 5; DB 4; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ALGWR 11
Db 50 ALGWR 54

RESULT 52
US-09-621-976-3972
; Sequence 3972, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
```

; SOFTWARE: Patent.pm
; SEQ ID NO 3972
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -36...-1
; NAME/KEY: UNSURE
; LOCATION: 7
; OTHER INFORMATION: Xaa = Arg,Thr
; NAME/KEY: UNSURE
; LOCATION: 24
; OTHER INFORMATION: Xaa = Asn,Thr
; NAME/KEY: UNSURE
; LOCATION: 30
; OTHER INFORMATION: Xaa = Ser,Thr
US-09-621-976-3972

Query Match 25.0%; Score 5; DB 4; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SSWAL 8
| | | | |
DB 45 SSWAL 49

RESULT 53
US-09-673-809-25
; Sequence 25, Application US/09673809
; Patent No. 6528261
; GENERAL INFORMATION:
; APPLICANT: INNOGENETICS N.V.
; TITLE OF INVENTION: Method for typing of HLA alleles.
; FILE REFERENCE: PCT99.86.HLA
; CURRENT APPLICATION NUMBER: US/09/673,809
; CURRENT FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 98870088.6
; PRIOR FILING DATE: 1998-04-20
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-673-809-25

Query Match 25.0%; Score 5; DB 4; Length 92;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRY 16
| | | | |
DB 77 WLRRY 81

RESULT 54
US-09-652-345-4
; Sequence 4, Application US/09652345
; Patent No. 6623933
; GENERAL INFORMATION:
; APPLICANT: Farb, David H
; APPLICANT: Russek, Shelley
; APPLICANT: Jang, Ming-Kuei
; APPLICANT: Gibbs, Terrell
; TITLE OF INVENTION: EFFECT OF STEROIDS ON NMDA RECEPTORS DEPENDS ON SUBUNIT
; TITLE OF INVENTION: COMPOSITION
; FILE REFERENCE: 0146-2026
; CURRENT APPLICATION NUMBER: US/09/652,345
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/151,802
; PRIOR FILING DATE: 1999-08-31

; PRIOR APPLICATION NUMBER: 09/378,547
; PRIOR FILING DATE: 1999-08-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-652-345-4

Query Match 25.0%; Score 5; DB 4; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ALGWR 11
| | | | |
DB 75 ALGWR 79

RESULT 55
US-09-540-236-2144
; Sequence 2144, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2144
; LENGTH: 98
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2144

Query Match 25.0%; Score 5; DB 4; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 WALGW 10
| | | | |
DB 11 WALGW 15

RESULT 56
US-09-710-279-1724
; Sequence 1724, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUB480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1724
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-1724

Query Match 25.0%; Score 5; DB 4; Length 108;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      15 RYCGW 19
      |||||
Db      97 RYCGW 101

RESULT 57
US-09-621-976-7483
; Sequence 7483, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7483
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-7483

Query Match      25.0%; Score 5; DB 4; Length 113;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QVWSS 5
      |||||
Db      5 QVWSS 9

RESULT 58
US-09-621-976-5186
; Sequence 5186, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5186
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-5186

Query Match      25.0%; Score 5; DB 4; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      13 LRRYG 17
      |||||
Db      17 LRRYG 21

RESULT 59
US-09-621-976-5890
; Sequence 5890, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
```

```
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5890
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -48..-1
US-09-621-976-5890

Query Match      25.0%; Score 5; DB 4; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 SSWAL 8
      |||||
Db      65 SSWAL 69

RESULT 60
US-09-513-999C-5792
; Sequence 5792, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5792
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 106
; OTHER INFORMATION: Xaa= * or Glu or Lys or Gln
US-09-513-999C-5792

Query Match      25.0%; Score 5; DB 4; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 VWSSW 6
      |||||
Db      90 VWSSW 94

RESULT 61
US-09-816-248-2
; Sequence 2, Application US/09816248
; Patent No. 6753411
; GENERAL INFORMATION:
; APPLICANT: BAUMANN, PETER
; APPLICANT: CECH, THOMAS R.
; TITLE OF INVENTION: PROTECTION-OF-TELOMERE-1 (POT-1) PROTEIN AND ENCODING
; FILE REFERENCE: POLYNUCLEOTIDES
; FILE REFERENCE: 089491/0201
; CURRENT APPLICATION NUMBER: US/09/816,248
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
```

```
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Stylonychia mytilis
US-09-816-248-2

Query Match      25.0%; Score 5; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 SSWAL 8
Db      117 SSWAL 121

RESULT 62
US-09-816-248-3
; Sequence 3, Application US/09816248
; Patent No. 6753411
; GENERAL INFORMATION:
; APPLICANT: BAUMANN, PETER
; APPLICANT: CECH, THOMAS R.
; TITLE OF INVENTION: PROTECTION-OF-TELOMERE-1 (POT-1) PROTEIN AND ENCODING
; TITLE OF INVENTION: POLYNUCLEOTIDES
; FILE REFERENCE: 089491/0201
; CURRENT APPLICATION NUMBER: US/09/816,248
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Oxytricha trifallax
US-09-816-248-3

Query Match      25.0%; Score 5; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 SSWAL 8
Db      117 SSWAL 121

RESULT 63
US-09-816-248-4
; Sequence 4, Application US/09816248
; Patent No. 6753411
; GENERAL INFORMATION:
; APPLICANT: BAUMANN, PETER
; APPLICANT: CECH, THOMAS R.
; TITLE OF INVENTION: PROTECTION-OF-TELOMERE-1 (POT-1) PROTEIN AND ENCODING
; TITLE OF INVENTION: POLYNUCLEOTIDES
; FILE REFERENCE: 089491/0201
; CURRENT APPLICATION NUMBER: US/09/816,248
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Oxytricha nova
US-09-816-248-4

Query Match      25.0%; Score 5; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 SSWAL 8
Db      117 SSWAL 121

RESULT 64
US-09-816-248-5
; Sequence 5, Application US/09816248
; Patent No. 6753411
; GENERAL INFORMATION:
; APPLICANT: BAUMANN, PETER
; APPLICANT: CECH, THOMAS R.
; TITLE OF INVENTION: PROTECTION-OF-TELOMERE-1 (POT-1) PROTEIN AND ENCODING
; TITLE OF INVENTION: POLYNUCLEOTIDES
; FILE REFERENCE: 089491/0201
; CURRENT APPLICATION NUMBER: US/09/816,248
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Oxytricha nova
US-09-816-248-5

Query Match      25.0%; Score 5; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 SSWAL 8
Db      117 SSWAL 121

RESULT 65
US-09-540-236-3251
; Sequence 3251, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: GARY L. BRETON ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARACTAE
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3251
```

```
US-08-980-523-11
; Sequence 11, Application US/08980523
; Patent No. 6310181
; GENERAL INFORMATION:
; APPLICANT: KOUHARA, HARUHIKO
; APPLICANT: SPIVAK-KROIZMAN, Taly
; APPLICANT: LAX, Irit
; APPLICANT: SCHLESSINGER, JOSEPH
; TITLE OF INVENTION: ADAPTOR PROTEIN FR52 AND
; TITLE OF INVENTION: RELATED PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/980,523
; FILING DATE: December 1, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/21851
; FILING DATE: December 1, 1997
; APPLICATION NUMBER: 60/032,093
; FILING DATE: December 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 230/045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-980-523-11

Query Match      25.0%; Score 5; DB 3; Length 129;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      13 LRRYG 17
Db      52 LRRYG 56

RESULT 65
US-09-540-236-3251
; Sequence 3251, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: GARY L. BRETON ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARACTAE
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3251
```



```
; LENGTH: 131
; TYPE: PRT
; ORGANISM: M.cattarrhalis
US-09-540-236-3251

Query Match      25.0%; Score 5; DB 4; Length 131;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 SSWAL 8
Db      59 SSWAL 63

RESULT 66
US-09-036-574-4
; Sequence 4, Application US/09036574
; Patent No. 6215045
; GENERAL INFORMATION:
; APPLICANT: (countries other than US) THE UNIVERSITY OF
; APPLICANT: MELBOURNE
; APPLICANT: (US only) KNOX, RB; SINGH, MB; and XU, H.
; TITLE OF INVENTION: DEVELOPMENTAL REGULATION IN ANOTHER TISSUE OF PLANTS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/454,294A
; FILING DATE: 16-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PCT INTERNATIONAL
; FILING DATE: 15-DEC-1993
; APPLICATION NUMBER: AU PL6400
; FILING DATE: 16-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLIO, FRANK S
; REFERENCE/DOCKET NUMBER: 9738
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: + 1 516 742 4343
; TELEFAX: + 1 516 742 4366
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-454-294A-4

Query Match      25.0%; Score 5; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 SSWAL 8
Db      119 SSWAL 123

RESULT 68
US-09-270-767-32020
; Sequence 32020, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32020
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32020

Query Match      25.0%; Score 5; DB 4; Length 144;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      11 RWLR 15
```

```
Db          62 RWLR 66
|||||
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20364
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20364

Query Match      25.0%; Score 5; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      10 RWLR 14
|||||
Db      90 RWLR 94

RESULT 72
US-09-252-991A-30988
; Sequence 30988, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30988
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30988

Query Match      25.0%; Score 5; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      10 RWLR 14
|||||
Db      43 RWLR 47

RESULT 73
US-09-252-991A-32019
; Sequence 32019, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32019
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32019
```

```
Db          62 RWLR 66
|||||
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 47237
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47237

Query Match      25.0%; Score 5; DB 4; Length 144;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      11 RWLR 15
|||||
Db      62 RWLR 66

RESULT 70
US-09-252-991A-27239
; Sequence 27239, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27239
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27239

Query Match      25.0%; Score 5; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      11 RWLR 15
|||||
Db      21 RWLR 25

RESULT 71
US-09-252-991A-20364
; Sequence 20364, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

```
Query Match      25.0%; Score 5; DB 4; Length 169;
Best Local Similarity 100.0%; Pred.No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      14 RRYGW 18
Db      88 RRYGW 92

RESULT 74
US-09-252-991A-18579
; Sequence 18579, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18579
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18579

Query Match      25.0%; Score 5; DB 4; Length 170;
Best Local Similarity 100.0%; Pred.No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      11 RWLRR 15
Db      105 RWLRR 109

RESULT 75
US-09-252-991A-20339
; Sequence 20339, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20339
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20339

Query Match      25.0%; Score 5; DB 4; Length 173;
Best Local Similarity 100.0%; Pred.No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      11 RWLRR 15
Db      24 RWLRR 28

Search completed: October 26, 2004, 07:25:04
Job time : 22.25 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 07:05:46 ; Search time 16.5 Seconds

(without alignments)
116.626 Million cell updates/sec

Title: US-10-066-965A-3

Perfect score: 20

Sequence: 1 PRGAPMWRWVCQMLETMFL 20

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : PIR_79:*

1: Pirl:*

2: Pirl:*

3: Pirl:*

4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	30.0	191	A55228	fms protein homolo
2	6	30.0	232	H75274	polypeptide deform
3	6	30.0	352	S17313	transcription fact
4	6	30.0	377	A34672	transcription fact
5	6	30.0	1437	S07430	M polypeptide prec
6	5	25.0	34	PQ0341	L protein - rabies
7	5	25.0	34	PQ0363	L protein - rabies
8	5	25.0	75	T26778	hypothetical prote
9	5	25.0	77	F69219	conserved hypothet
10	5	25.0	99	B82524	hypothetical prote
11	5	25.0	110	T17913	hypothetical prote
12	5	25.0	119	T08271	probable thioredox
13	5	25.0	123	D72579	hypothetical prote
14	5	25.0	128	I12372	spasmolytic protei
15	5	25.0	129	I53637	spasmolytic polype
16	5	25.0	140	T33359	hypothetical prote
17	5	25.0	142	G82796	hypothetical prote
18	5	25.0	148	AI1645	hypothetical prote
19	5	25.0	148	AD1594	hypothetical prote
20	5	25.0	162	D75295	DNA topology modul
21	5	25.0	164	A97402	probable N-acetyl
22	5	25.0	164	AI2619	acetyltransferase
23	5	25.0	169	A82373	polypeptide deform
24	5	25.0	170	AF2621	polypeptide deform
25	5	25.0	170	F97403	polypeptide deform
26	5	25.0	173	G87282	polypeptide deform
27	5	25.0	175	B71732	methionyl-tRNA def
28	5	25.0	179	S58124	hypothetical prote
29	5	25.0	189	S39075	auxin-induced prot

30	5	25.0	195	1	SAVLDM	delta large antige
31	5	25.0	195	1	SAVLDS	delta large antige
32	5	25.0	196	2	T10941	auxin-induced prot
33	5	25.0	202	2	I46688	complement compone
34	5	25.0	209	2	C83079	hypothetical prote
35	5	25.0	220	2	B81014	conserved hypothet
36	5	25.0	220	2	A81959	probable pseudourid
37	5	25.0	222	2	S33204	hypothetical prote
38	5	25.0	222	2	S37734	hypothetical prote
39	5	25.0	230	2	S18997	hemagglutinin asso
40	5	25.0	230	2	C98274	nikp2 protein (A22
41	5	25.0	254	2	C98274	flagellar protein
42	5	25.0	258	2	H83954	lacto-N-neotetraos
43	5	25.0	268	2	B81971	lacto-N-neotetraos
44	5	25.0	275	2	C81027	glycosyl transfera
45	5	25.0	275	2	S70814	glycosyl transfera
46	5	25.0	276	2	S70815	endonuclease iv (e
47	5	25.0	277	2	E90589	hypothetical prote
48	5	25.0	278	2	T27610	lacto-N-neotetraos
49	5	25.0	279	2	A81971	lacto-N-neotetraos
50	5	25.0	280	2	A81027	hypothetical prote
51	5	25.0	280	2	E70745	hypothetical prote
52	5	25.0	284	2	T22501	hypothetical prote
53	5	25.0	288	2	A81529	transcription regu
54	5	25.0	294	2	C72378	sugar ABC transpor
55	5	25.0	296	2	T24827	hypothetical prote
56	5	25.0	304	2	G72776	hypothetical prote
57	5	25.0	305	2	T20906	hypothetical prote
58	5	25.0	308	2	A98986	malonyl CoA-acyl c
59	5	25.0	308	2	H75049	hypothetical prote
60	5	25.0	320	2	G71110	hypothetical prote
61	5	25.0	325	2	AB2574	calpain (EC 3.4.22
62	5	25.0	326	2	H82455	IS5 transposase VC
63	5	25.0	330	2	T02347	probable lipid tra
64	5	25.0	344	1	A40005	hyoscyanine (6S)-d
65	5	25.0	351	2	D96761	unknown protein (i
66	5	25.0	355	2	T29932	hypothetical prote
67	5	25.0	357	2	T24137	hypothetical prote
68	5	25.0	364	2	T08903	hypothetical prote
69	5	25.0	368	2	T40115	uv excision repair
70	5	25.0	382	2	D82264	probable exopolyas
71	5	25.0	394	1	BWH1XD	bexd protein - Hae
72	5	25.0	406	2	S39965	hypothetical prote
73	5	25.0	407	2	T00693	hypothetical prote
74	5	25.0	409	2	T24138	hypothetical prote
75	5	25.0	420	2	D87492	hypothetical prote
76	5	25.0	433	2	T41038	atp-specific succi
77	5	25.0	433	2	T50395	actin-related prot
78	5	25.0	435	2	G84518	Mutator-like trans
79	5	25.0	443	2	AB2841	MFS permease[proli
80	5	25.0	443	2	D97618	probable mfs trans
81	5	25.0	445	2	C83674	phenylacetyl-CoA l
82	5	25.0	446	2	T34782	probable signal pe
83	5	25.0	454	2	E95941	probable guanine d
84	5	25.0	454	2	G82364	DNA-damage-inducib
85	5	25.0	455	2	JC1224	nucleobindin precu
86	5	25.0	458	2	T31631	hypothetical prote
87	5	25.0	459	2	I55472	Calcium binding pr
88	5	25.0	469	2	H81986	probable peptidase
89	5	25.0	469	2	C81042	penicillin-binding
90	5	25.0	481	2	C82421	conserved hypothet
91	5	25.0	485	2	D84475	probable Athlia re
92	5	25.0	488	2	A27353	collagen alpha 1(I
93	5	25.0	492	2	T16659	hypothetical prote
94	5	25.0	501	2	G95103	pyruvate kinase fi
95	5	25.0	501	2	E97971	pyruvate kinase (E
96	5	25.0	511	2	JC7682	spermatogenesis as
97	5	25.0	516	2	T37066	probable integrat
98	5	25.0	527	2	E75503	probable fibronect
99	5	25.0	530	2	G96588	hypothetical prote
100	5	25.0	534	2	S21961	proline-rich prote

ALIGNMENTS

```
RESULT 1
fms protein homolog - Thermus aquaticus (fragment)
A:Species: Thermus aquaticus
C:Date: 25-Aug-1995 #sequence_revision 25-Aug-1995 #text_change 15-Mar-2004
C:Accession: A55228
R:Meinzel, T.; Blanquet, S.
J. Bacteriol. 176, 7387-7390, 1994
A:Title: Characterization of the Thermus thermophilus locus encoding peptide deformylase
A:Reference number: A55228; MUID:95050326; PMID:7961514
A:Accession: A55228
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-191 <MEI>
A:Cross-references: GB:X79087; NID:g602912; PIDN:CAA55695.1; PID:g602914
C:Superfamily: peptide deformylase

Query Match      30.0%; Score 6; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 MLETMF 19
Db 32 MLETMF 37

RESULT 2
H75274
polypeptide deformylase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: H75274
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: H75274
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-232 <WHI>
A:Cross-references: UNIPROT:Q9RRQ4; GB:AE002073; GB:AE000513; NID:g6460244; PIDN:AAFI197
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2434
A:Map position: 1
C:Superfamily: peptide deformylase

Query Match      30.0%; Score 6; DB 2; Length 232;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 MLETMF 19
Db 62 MLETMF 67

RESULT 3
S17313
transcription factor Oct-3, short splice form - mouse
N:Alternate names: NF-A3; transcription factor Oct-4
C:Species: Mus musculus (house mouse)
C:Date: 22-Nov-1993 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C:Accession: S17313; S10924; J34206
R:Okazawa, H.; Okamoto, K.; Ishino, F.; Ishino-Kaneko, T.; Takeda, S.; Toyoda, Y.; Muram
EMBO J. 10, 2997-3005, 1991
A:Title: The oct3 gene, a gene for an embryonic transcription factor, is controlled by a
A:Reference number: S17313; MUID:92007744; PMID:1915274
A:Accession: S17313
```

```
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-352 <OKA>
A:Cross-references: UNIPROT:P20263; GB:S58426; NID:g235987; PIDN:AAI9896.1; PID:g235988
R:Rosner, M.H.; Viganò, M.A.; Ozato, K.; Timmons, P.M.; Poirier, F.; Rigby, P.W.J.; Stau
Nature 345, 686-692, 1990
A:Title: A POU-domain transcription factor in early stem cells and germ cells of the man
A:Reference number: S10924; MUID:90294899; PMID:1972777
A:Accession: S10924
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-30, 'P', 32-352 <ROS>
R:Schoeller, H.R.; Ruppert, S.; Suzuki, N.; Chowdhury, K.; Gruss, P.
Nature 344, 435-439, 1990
A:Title: New type of POU domain in germ line-specific protein Oct-4.
A:Reference number: A34206; MUID:90206045; PMID:1690859
A:Accession: A34206
A:Molecule type: mRNA
A:Residues: 'M', 30, 'P', 32-352 <SCH>
A:Cross-references: GB:X52437; NID:g53500; PIDN:CAA36682.1; PID:g53501
C:Superfamily: transcription factor Oct-3; homeobox homology; POU domain homology
C:Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulati
F:138-205/Domain: POU domain homology <POU>
F:224-280/Domain: homeobox homology <HOX>

Query Match      30.0%; Score 6; DB 2; Length 352;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LETMFL 20
Db 238 LETMFL 243

RESULT 4
A34672
transcription factor Oct-3, long splice form - mouse
N:Alternate names: NF-A3; transcription factor Oct-4
C:Species: Mus musculus (house mouse)
C:Date: 13-Jul-1990 #sequence_revision 09-Oct-1992 #text_change 09-Jul-2004
C:Accession: A34672
R:Okamoto, K.; Okazawa, H.; Okuda, A.; Sakai, M.; Muramatsu, M.; Hamada, H.
Cell 60, 461-472, 1990
A:Title: A novel octamer binding transcription factor is differentially expressed in mou
A:Reference number: A34672; MUID:90150273; PMID:1967980
A:Accession: A34672
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-377 <OKA>
A:Cross-references: UNIPROT:P20263; GB:M34381; NID:g200117; PIDN:AAA39844.1; PID:g200118
A:Note: the authors translated the codon GAT for residue 19 as Tyr
C:Superfamily: transcription factor Oct-3; homeobox homology; POU domain homology
C:Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulati
F:138-205/Domain: POU domain homology <POU>
F:224-280/Domain: homeobox homology <HOX>

Query Match      30.0%; Score 6; DB 2; Length 377;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LETMFL 20
Db 238 LETMFL 243

RESULT 5
S07430
M polypeptide precursor - Germiston virus
N:Contains: glycoprotein G1; glycoprotein G2; nonstructural protein NS-M
C:Species: Germiston virus
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: S07430
R:Pardigon, N.; Vialat, P.; Gerbaud, S.; Girard, M.; Bouloy, M.
```

Virus Res. 11, 73-85, 1988
 A:Title: Nucleotide sequence of the M segment of Germiston virus: comparison of the M gene
 A:Reference number: S07430; MUID:89021381; PMID:3176688
 A:Accession: S07430
 A:Molecule type: genomic RNA
 A:Residues: 1-1437 <PAR>
 A:Cross-references: UNIPROT:P12430; EMBL:M21951
 C:Genetics:
 A:Gene: M
 C:Superfamily: bunyavirus M polypeptide
 C:Keywords: glycoprotein; nonstructural protein; polypeptide; transmembrane protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-1437/Product: M polypeptide #status predicted <MPP>
 F:1382-1412/Domain: transmembrane #status predicted <TMM>
 F:65,88,252,474,627,897,1173/Binding site: carbohydrate (Asn) (covalent) #status predicted
 Query Match 30.0%; Score 6; DB 2; Length 1437;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 15 LETMFL 20
 Db 560 LETMFL 565
 |||||
 RESULT 6
 PQ0341
 L protein - rabies virus
 C:Species: rabies virus
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C:Accession: PQ0341
 R:Sacramento, D.; Badrane, H.; Bourhy, H.; Tordo, N.
 J. Gen. Virol. 73, 1149-1158, 1992
 A:Title: Molecular epidemiology of rabies virus in France: Comparison with vaccine strain
 A:Reference number: PQ0340; MUID:92268873; PMID:1588319
 A:Accession: PQ0341
 A:Molecule type: mRNA
 A:Residues: 1-34 <SAC>
 A:Cross-references: UNIPROT:Q9IPJ5
 A:Note: this sequence was obtained from isolates: WR56, WR06, WR39, WR97, WR18, WR63, WR
 C:Superfamily: rhabdovirus L protein
 Query Match 25.0%; Score 5; DB 2; Length 34;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PRGAP 5
 Db 21 PRGAP 25
 |||||
 RESULT 7
 PQ0363
 L protein - rabies virus (isolate WR17)
 C:Species: rabies virus
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C:Accession: PQ0363
 R:Sacramento, D.; Badrane, H.; Bourhy, H.; Tordo, N.
 J. Gen. Virol. 73, 1149-1158, 1992
 A:Title: Molecular epidemiology of rabies virus in France: Comparison with vaccine strain
 A:Reference number: PQ0340; MUID:92268873; PMID:1588319
 A:Accession: PQ0363
 A:Molecule type: mRNA
 A:Residues: 1-34 <SAC>
 A:Cross-references: UNIPROT:Q9IPJ5
 C:Superfamily: rhabdovirus L protein
 Query Match 25.0%; Score 5; DB 2; Length 34;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PRGAP 5
 Db 21 PRGAP 25
 |||||

Db 21 PRGAP 25
 RESULT 8
 T26778
 hypothetical protein Y40B1A.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T26778
 R:Harris, B.
 submitted to the EMBL Data Library, December 1998
 A:Reference number: Z20264
 A:Accession: T26778
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-75 <MIL>
 A:Cross-references: UNIPROT:Q9XW29; EMBL:AL034392; PIDN:CAA22304.1; CBSP:Y40B1A.1
 A:Experimental source: clone Y40B1A
 C:Genetics:
 A:Gene: CBSP:Y40B1A.1
 A:Introns: 61/3
 Query Match 25.0%; Score 5; DB 2; Length 75;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PRGAP 5
 Db 5 PRGAP 9
 |||||
 RESULT 9
 F69219
 conserved hypothetical protein MTH895 - Methanobacterium thermoautotrophicum (strain Delt
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: F69219
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; F
 ; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: F69219
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-77 <MTH>
 A:Cross-references: UNIPROT:O26981; GB:AE000865; GB:AE000666; NID:g2621984; PIDN:AAB8539;
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH895
 C:Superfamily: probable glutaredoxin grx-2
 Query Match 25.0%; Score 5; DB 2; Length 77;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 CQMLE 16
 Db 14 CQMLE 18
 |||||
 RESULT 10
 B82524
 hypothetical protein XF2711 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C:Accession: B82524
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below

A:Accession: B82524
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-99 <SIM>
 A:Cross-references: UNIPROT:Q9PA09; GB:AE004077; GB:AE003849; NID:g9107952; PIDN:AAF8550
 A:Experimental source: strain 9a5c
 R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Bionesi, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Fromm J.D.; Junqueira, M.B.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF2711

Query Match 25.0%; Score 5; DB 2; Length 99;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETM 18
 |||||
 Db 1 MLETM 5

RESULT 11
 T17913
 hypothetical protein A410L - Chlorella virus PBCV-1
 C:Species: Chlorella virus PBCV-1
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T17913
 R:Graves, M.V.; Van Etten, J.L. submitted to the EMBL Data Library, May 1999
 A:Reference number: Z18806
 A:Accession: T17913
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-110 <GRA>
 A:Cross-references: UNIPROT:Q98462; EMBL:U42580; NID:g4028896; PIDN:AAC96778.1
 A:Experimental source: specific host Chlorella strain NC64
 C:Genetics:
 A:Note: A410L

Query Match 25.0%; Score 5; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETMFL 20
 |||||
 Db 67 ETMFL 71

RESULT 12
 T08271
 probable thioredoxin - Halobacterium sp. (strain NRC-1) plasmid pNRC100
 N:Alternate names: hypothetical protein H0606; hypothetical protein H1757
 C:Species: Halobacterium sp.
 A:Variety: strain NRC-1
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
 C:Accession: T08271; T08372
 R:Ng, W.V.; Ciufo, S.A.; Smith, T.M.; Bumgarner, R.E.; Baskin, D.; Faust, J.; Hall, B.; Genome Res. 8, 1131-1141, 1998
 A:Title: Snapshot of a large dynamic replicon in a halophilic Archaeon: megaplasmid or A:Reference number: Z16408; MUID:99063795; PMID:9847077
 A:Accession: T08271
 A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-119 <NGW>
 A:Cross-references: UNIPROT:O46709; EMBL:AF016485; NID:g2822278; PID:g2822332; HALOSP:H1757
 A:Experimental source: strain NRC-1
 A:Genetics: COP1
 A:Accession: T08372
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-119 <DAS>
 A:Cross-references: EMBL:AF016485; NID:g2822278; PID:g2822433; HALOSP:H1757
 A:Experimental source: strain NRC-1
 A:Genetics: COP2
 C:Genetics: <COP1>
 A:Gene: trxA; HALOSP:H0606
 A:Genome: plasmid pNRC100
 C:Genetics: <COP2>
 A:Gene: trxA; HALOSP:H1757
 A:Genome: plasmid pNRC100
 C:Function:
 A:Description: involved in reduction of ribonucleotides, methionine sulfoxide sulfate, A:Superfamily: thioredoxin; thioredoxin homology

Query Match 25.0%; Score 5; DB 2; Length 119;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 COMLE 16
 |||||
 Db 48 COMLE 52

RESULT 13
 D72579
 hypothetical protein APE1916 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 C:Accession: D72579
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: D72579
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-123 <KAW>
 A:Cross-references: UNIPROT:Q9YAM7; DDBJ:AP000062; NID:g5105244; PIDN:BAA80921.1; PID:d A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE1916

Query Match 25.0%; Score 5; DB 2; Length 123;
 Best Local Similarity 100.0%; Pred. No. 81;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 |||||
 Db 34 PRGAP 38

RESULT 14
 S12372
 spasmodytic protein precursor - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 21-Nov-1993 #sequence_revision 24-May-1996 #text_change 18-Jun-1999
 C:Accession: S12372; S63997
 R:Tomasetto, C.; Rio, M.C.; Gautier, C.; Wolf, C.; Hareuveni, M.; Chambon, P.; Lathe, R.; EWO J. 9, 407-414, 1990
 A:Title: hsp, the domain-duplicated homolog of p52 protein, is co-expressed with p52 in A:Reference number: S12371; MUID:90151615; PMID:2303034
 A:Accession: S12372
 A:Molecule type: mRNA
 A:Residues: 1-128 <TOM>

A;Cross-references: EMBL:X51697
 R;Tomasatto, C.; Rio, M.C.; Gautier, C.; Wolf, C.; Hareuveni, M.; Chambon, P.; Lathe, R. submitted to the EMBL Data Library, December 1993
 A;Description: hsp, the domain-duplicated homolog of p52 protein, is co-expressed with pA;Reference number: S63997
 A;Accession: S63997
 A;Molecule type: mRNA
 A;Residues: 1-10, 'V', 12-128 <TOW>
 A;Cross-references: EMBL:X51697; NID:G54165; PIDN:CAA35994.1; PID:G54166
 C;Genetics: msp
 C;Function: hsp
 A;Description: inhibits gastrointestinal motility and gastric acid secretion
 C;Superfamily: spasmodic protein; trefoil homology
 C;Keywords: duplication; hormone; pancreas
 F;1-22/Domain: signal sequence (fragment) #status predicted <SIG>
 F;23-128/Product: spasmodic protein (fragment) #status predicted <MAT>
 F;30-71/Domain: trefoil homology <TRF1>
 F;80-120/Domain: trefoil homology <TRF2>
 F;128-126,30-57,41-56,51-68,80-106,90-105,100-117/Disulfide bonds: #status predicted

Query Match 25.0%; Score 5; DB 1; Length 128;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 |||||
 Db 2 PRGAP 6

RESULT 15
 153637
 spasmodic polypeptide - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C;Accession: I53637
 R;Jeffrey, G.P.; Oates, P.S.; Wang, T.C.; Babyatsky, M.W.; Brand, S.J.
 Gastroenterology 105, 336-345, 1994
 A;Title: Spasmodic polypeptide: a trefoil peptide secreted by rat gastric mucous cells
 A;Reference number: I53637; PMID:94131228; PMID:8299900
 A;Accession: I53637
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-129 <RES>
 A;Cross-references: UNIPROT:Q09030; GB:M97255; NID:G207049; PIDN:AAA19025.1; PID:G207050
 C;Superfamily: spasmodic protein; trefoil homology
 F;31-72/Domain: trefoil homology <TRF1>
 F;81-121/Domain: trefoil homology <TRF2>

Query Match 25.0%; Score 5; DB 1; Length 129;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 |||||
 Db 3 PRGAP 7

RESULT 16
 T33359
 hypothetical protein F16G10.7 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T33359
 R;Gattung, S.; Scheet, P.
 submitted to the EMBL Data Library, July 1998
 A;Description: The sequence of C. elegans cosmid F16G10.
 A;Reference number: 221329
 A;Accession: T33359
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-140 <GAT>
 A;Cross-references: UNIPROT:O76592; EMBL:AF077537; PIDN:AAC26275.1; GSPDB:GN00020; CBSP:

A;Experimental source: strain Bristol N2; clone F16G10
 C;Genetics:
 A;Gene: CBSP:F16G10.7
 A;Map position: 2
 A;Introns: 12/1; 31/1; 75/3

Query Match 25.0%; Score 5; DB 2; Length 140;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
 |||||
 Db 25 LETMF 29

RESULT 17
 G82796
 hypothetical protein XF0524 [imported] - Xylella fastidiosa (strain 9a5c)
 C;Species: Xylella fastidiosa
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C;Accession: G82796
 R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
 Nature 406, 151-157, 2000
 A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A;Reference number: A82515; MUID:20365717; PMID:10910347
 A;Note: for a complete list of authors see reference number A59328 below
 A;Accession: G82796
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-142 <SIM>
 A;Cross-references: UNIPROT:Q9PFY2; GB:AE003900; GB:AE003849; NID:G9105366; PIDN:AAF83334
 A;Experimental source: strain 9a5c
 R;Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al-Briteson, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000
 A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigret, Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.; A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Mitacca, B.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki, M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

Query Match 25.0%; Score 5; DB 2; Length 142;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 |||||
 Db 59 PRGAP 63

RESULT 18
 AH1645
 hypothetical protein lin1705 [imported] - Listeria innocua (strain Clip11262)
 C;Species: Listeria innocua
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C;Accession: AH1645
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: AH1645

A;Status: preliminary
A:Molecule type: DNA
A;Residues: 1-148 <GLA>
A;Cross-references: UNIPROT:Q92B45; GB:AL592022; PIDN:CAC96936.1; PID:G16414192; GSPDB:G
A;Experimental source: strain Clp11262
C;Genetics:
A;Gene: lin1705

Query Match 25.0%; Score 5; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
Db 11 QMLET 15

RESULT 19
AD1594
hypothetical protein lin1293 [imported] - *Listeria innocua* (strain Clp11262)
C;Species: *Listeria innocua*
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AD1594
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Roland, J.A.; Voss, H.; Wehlend,
A;Title: Comparative genomics of *Listeria* species
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AD1594
A;Status: preliminary
A:Molecule type: DNA
A;Residues: 1-148 <GLA>
A;Cross-references: UNIPROT:Q92C98; GB:AL592022; PIDN:CAC96524.1; PID:G16413766; GSPDB:G
A;Experimental source: strain Clp11262
C;Genetics:
A;Gene: lin1293

Query Match 25.0%; Score 5; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
Db 11 QMLET 15

RESULT 20
D75295
DNA topology modulation protein Flar-related protein - *Deinococcus radiodurans* (strain R
C;Species: *Deinococcus radiodurans*
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: D75295
R;White, O.; Elsen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.N.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: D75295
A;Status: preliminary
A:Molecule type: DNA
A;Residues: 1-162 <WHI>
A;Cross-references: UNIPROT:Q9RS54; GB:AE002059; GB:AE000513; NID:G6460070; PIDN:AAF1181
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2273
A;Map position: 1

Query Match 25.0%; Score 5; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 1e+02;

QY 13 QMLET 17
Db 11 QMLET 15

us-10-066-965a-3.oligo.rpr

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
Db 133 QMLET 137

RESULT 21
A97402
Probable N-acetyltransferase (AF232919) [imported] - *Agrobacterium tumefaciens* (strain C
C;Species: *Agrobacterium tumefaciens*
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: A97402
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ourullo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: A97402
A;Status: preliminary
A:Molecule type: DNA
A;Residues: 1-164 <KUR>
A;Cross-references: UNIPROT:Q8UIE4; GB:AE007869; PIDN:AAK86170.1; PID:G15155263; GSPDB:G
C;Genetics:
A;Gene: AGR_C 619
A;Map position: circular chromosome

Query Match 25.0%; Score 5; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETMFL 20
Db 114 ETMFL 118

RESULT 22
AI2619
acetyltransferase [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
C;Species: *Agrobacterium tumefaciens*
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AI2619
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Monks, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AI2619
A;Status: preliminary
A:Molecule type: DNA
A;Residues: 1-164 <KUR>
A;Cross-references: UNIPROT:Q8UIE4; GB:AE008688; PIDN:AAL41375.1; PID:G17738691; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu0353
A;Map position: circular chromosome

Query Match 25.0%; Score 5; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETMFL 20
Db 114 ETMFL 118

RESULT 23
A82373
polypeptide deformylase VC0046 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
C;Species: *Vibrio cholerae*

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82373
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: A82373
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-169 <HEI>
A;Cross-references: UNIPROT:Q9KVU3; GB:AE004096; GB:AE003852; NID:g9654440; PIDN:AAF9322
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0046
A;Map position: 1
C;Superfamily: peptide deformylase

Query Match 25.0%; Score 5; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0

QY 14 MLETM 18
|||||
Db 35 MLETM 39

RESULT 24
AF2621
polypeptide deformylase def [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AF2621
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, B.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AF2621
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-170 <KUR>
A;Cross-references: UNIPROT:Q8UID1; GB:AE008688; PIDN:AAI41388.1; PID:gl7738706; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: def
A;Map position: circular chromosome
C;Superfamily: peptide deformylase

Query Match 25.0%; Score 5; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETM 18
|||||
Db 35 MLETM 39

RESULT 25
F97403
polypeptide deformylase (AF213822) [imported] - Agrobacterium tumefaciens (strain C58, C
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: F97403
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Lib, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2233-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: F97403
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-170 <KUR>
A;Cross-references: UNIPROT:Q8UID1; GB:AE007869; PIDN:AAK86183.1; PID:gl5155278; GSPDB:G
C;Genetics:
A;Gene: AGR_C 640
A;Map position: circular chromosome
C;Superfamily: peptide deformylase

Query Match 25.0%; Score 5; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETM 18
|||||
Db 35 MLETM 39

RESULT 26
G87282
polypeptide deformylase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: G87282
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.I
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87282
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-173 <STO>
A;Cross-references: UNIPROT:Q9ABF5; GB:AE005673; NID:gl3421407; PIDN:AAK22259.1; GSPDB:G
C;Genetics:
A;Gene: CC0272
C;Superfamily: peptide deformylase

Query Match 25.0%; Score 5; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETM 18
|||||
Db 39 MLETM 43

RESULT 27
B71732
methionyl-tRNA deformylase (def) RP208 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C;Accession: B71732
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark, U.
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: B71732
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-175 <AND>
A;Cross-references: UNIPROT:Q92DV8; GB:AJ235270; GB:AJ235269; NID:g3860572; PIDN:CAA1467
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: def; RP208
C;Superfamily: peptide deformylase

Query Match 25.0%; Score 5; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 1.e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETM 18
|||||
DB 35 MLETM 39

RESULT 28

S58124
hypothetical protein - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C:Accession: S58124
R:Blaszczyk, R.; Loehiger, C.; Wehling, J.
submitted to the EMBL Data Library, July 1995
A:Reference number: S58124
A:Accession: S58124
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-179 <BLA>
A:Cross-references: UNIPROT:Q29839; EMBL:X89707; NID:g1052702; PIDN:CAA61854.1; PID:g105
C:Genetics:
A:Introns: 89/2

Query Match 25.0%; Score 5; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
DB 15 PRGAP 19

RESULT 29

S39075
auxin-induced protein IAA4/5 - garden pea
C:Species: Pisum sativum (garden pea)
C>Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C:Accession: S39075; S39077
R:Oeller, P.W.; Keller, J.A.; Parks, J.E.; Silbert, J.E.; Theologis, A.
J. Mol. Biol. 233, 789-798, 1993
A:Title: Structural characterization of the early indoleacetic acid-inducible genes, PS-
A:Reference number: S39075; MUID:94016594; PMID:8411182
A:Accession: S39075
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-189 <OEL>
A:Cross-references: UNIPROT:P49679; EMBL:X68216
A:Experimental source: var. Alaska
A:Accession: S39077
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-189 <OEW>
A:Cross-references: EMBL:X68215; NID:g414180; PIDN:CAA48297.1; PID:g414181
A:Experimental source: var. Alaska
C:Genetics:
A:Gene: IAA4/5
A:Introns: 68/2; 132/1; 165/2
C:Superfamily: auxin-induced protein aux28

Query Match 25.0%; Score 5; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMP 19
|||||
DB 123 LETMP 127

RESULT 30

SAVLDM
delta large antigen - hepatitis delta virus (strain Japanese M-1)
N:Alternate names: HDag
C:Species: hepatitis delta virus
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004

C:Accession: A36409
R:Imazeki, F.; Omata, M.; Ohto, M.
J. Virol. 64, 5594-5599, 1990
A:Title: Heterogeneity and evolution rates of delta virus RNA sequences.
A:Reference number: A36409; MUID:91012805; PMID:2214027
A:Accession: A36409
A:Molecule type: genomic RNA
A:Residues: 1-195 <IMA>
A:Cross-references: UNIPROT:P25881; GB:D90190; GB:M58299; NID:g221691; PIDN:BAA14214.1;
C:Superfamily: hepatitis delta virus large antigen
C:Keywords: core protein
F:2-195/Product: Delta large antigen #status predicted <MAT>

Query Match 25.0%; Score 5; DB 1; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
DB 159 PRGAP 163

RESULT 31

SAVLDS
delta large antigen - hepatitis delta virus (isolate Japanese S-1)
N:Alternate names: HDag
C:Species: hepatitis delta virus
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C:Accession: B36409; S18678
R:Imazeki, F.; Omata, M.; Ohto, M.
J. Virol. 64, 5594-5599, 1990
A:Title: Heterogeneity and evolution rates of delta virus RNA sequences.
A:Reference number: A36409; MUID:91012805; PMID:2214027
A:Accession: B36409
A:Molecule type: genomic RNA
A:Residues: 1-195 <IMA1>
A:Cross-references: UNIPROT:P25883; GB:D90192; GB:M58303; NID:g221695; PIDN:BAA14216.1;
A:Experimental source: isolate Japanese S-1
R:Imazeki, F.; Omata, M.; Ohto, M.
Nucleic Acids Res. 19, 5439, 1991
A:Title: Complete nucleotide sequence of hepatitis delta virus RNA in Japan.
A:Reference number: S18678; MUID:92020244; PMID:1923832
A:Accession: S18678
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-195 <IMA2>
A:Cross-references: EMBL:X60193; NID:g59497; PIDN:CAA42749.1; PID:g59498
A:Experimental source: isolate Japanese S-1
A:Note: note this sequence was submitted to the EMBL Data Library, June 1991
C:Superfamily: hepatitis delta virus large antigen
C:Keywords: core protein
F:2-195/Product: delta large antigen #status predicted <MAT>

Query Match 25.0%; Score 5; DB 1; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
DB 159 PRGAP 163

RESULT 32

T10941
auxin-induced protein Aux22 - mung bean
C:Species: Vigna radiata (mung bean)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T10941
R:Yamamoto, K.
submitted to the EMBL Data Library, February 1993
A:Reference number: Z17206
A:Accession: T10941
A>Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-196 <YAM>

A;Cross-references: UNIPROT:P32294; EMBL:D14413

A;Experimental source: hypocoetyl

C;Genetics:

A;Gene: arg4

C;Superfamily: auxin-induced protein aux28

Query Match 25.0%; Score 5; DB 2; Length 196;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19

|||||

Db 130 LETMF 134

RESULT 33

I46688

complement component C8 gamma subunit - rabbit

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 09-Jul-2004

C;Accession: I46688

R;White, R.V.; Kaufman, K.M.; Letson, C.S.; Platteborze, P.L.; Sodetz, J.M.

J. Immunol. 152, 2501-2509, 1994

A;Title: Characterization of rabbit complement component C8: Functional evidence for the

A;Reference number: I46688; MUID:94179833; PMID:7510745

A;Accession: I46688

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-202 <WHI>

A;Cross-references: UNIPROT:Q28679; GB:I26979; NID:G469064; PIDN:AAA31193.1; PID:G469065

C;Superfamily: lipocalin; lipocalin; lipocalin homology

F;42-188/Domain: lipocalin homology <LIP>

Query Match 25.0%; Score 5; DB 2; Length 202;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5

|||||

Db 26 PRGAP 30

RESULT 34

C83079

hypothetical protein PA4535 [imported] - Pseudomonas aeruginosa (strain PAO1)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C;Accession: C83079

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: C83079

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-209 <STO>

A;Cross-references: UNIPROT:Q9HVP2; GB:AE004867; GB:AE004091; NID:G9950769; PIDN:AAG0792

A;Experimental source: strain PAO1

C;Genetics:

A;Gene: PA4535

Query Match 25.0%; Score 5; DB 2; Length 209;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 RWVCQ 13

|||||

Db 31 RWVCQ 35

RESULT 35

E81014

conserved hypothetical protein NMB2018 [imported] - Neisseria meningitidis (strain MC58 #

C;Species: Neisseria meningitidis

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C;Accession: E81014

R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

rr, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Masignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver

A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: A81000; MUID:20175755; PMID:10710307

A;Accession: E81014

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-220 <TET>

A;Cross-references: UNIPROT:Q9JXJ0; GB:AE002552; GB:AE002098; NID:G7227279; PIDN:AAF42343

A;Experimental source: serogroup B, strain MC58

C;Genetics:

A;Gene: NMB2018

Query Match 25.0%; Score 5; DB 2; Length 220;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETMFL 20

|||||

Db 173 ETMFL 177

RESULT 36

A81959

probable pseudouridine synthase NMA0422 [imported] - Neisseria meningitidis (strain Z2491

C;Species: Neisseria meningitidis

C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004

C;Accession: A81959

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel]

; Holroyd, S.; Jageis, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A;Reference number: A81775; MUID:20222556; PMID:10761919

A;Accession: A81959

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-220 <PAR>

A;Cross-references: UNIPROT:Q9JWE2; GB:AL162753; GB:AL157959; NID:G7379120; PIDN:CAB83721

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: NMA0422

Query Match 25.0%; Score 5; DB 2; Length 220;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETMFL 20

|||||

Db 173 ETMFL 177

RESULT 37

S33204

hypothetical protein - long-stalked stitchwort

C;Species: Stellaria longipes (long-stalked stitchwort)

C;Date: 09-Jun-1994 #sequence_revision 12-May-1995 #text_change 09-Jul-2004

C;Accession: S33204

R;Zhang, X.H.

submitted to the EMBL Data Library, April 1993

A;Reference number: S33204

A;Accession: S33204

A;Status: preliminary

A;Molecule type: mRNA

A:Residues: 1-222 <ZHA>
A:Cross-references: UNIPROT:Q41348; EMBL:X71601; NID:g2971175; PID:g2971176
C:Superfamily: hypothetical protein H11647

Query Match 25.0%; Score 5; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 54 PRGAP 58

RESULT 38
S37734
hypothetical protein - Vibrio cholerae
C:Species: Vibrio cholerae
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S37734
R:Franzon, V.L.; Barker, A.; Manning, P.A.
Infect. Immun. 61, 3032-3037, 1993
A:Title: Nucleotide sequence encoding the mannose-fucose-resistant hemagglutinin of Vibrio cholerae
A:Reference number: S37730; MUID:93293328; PMID:8514410
A:Accession: S37734
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-230 <FRA>
A:Cross-references: UNIPROT:Q56638; EMBL:X64097
C:Superfamily: type II site-specific DNA-methyltransferase

Query Match 25.0%; Score 5; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETMFL 20
|||||
Db 101 ETMFL 105

RESULT 39
S18997
hemagglutinin associated protein VCA0447 [imported] - Vibrio cholerae (strain N16961 serotype O1)
C:Species: Vibrio cholerae
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S18997; H82460
R:van Dongen, W.M.A.M.; van Vlerken, M.M.A.; de Graaf, F.K.
submitted to the EMBL Data Library, September 1991
A:Description: Nucleotide sequence of a DNA fragment encoding a Vibrio cholerae haemagglutinin
A:Reference number: S18996
A:Accession: S18997
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-230 <VAN>
A:Cross-references: UNIPROT:Q56638; EMBL:X64387; NID:948355; PIDN:CAA45727.1; PID:948357
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: H82460
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-230 <HEI>
A:Cross-references: GB:AB004377; GB:AB003853; NID:g9657831; PIDN:AAF96353.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0447
A:Map position: 2
C:Superfamily: type II site-specific DNA-methyltransferase

Query Match 25.0%; Score 5; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETMFL 20
|||||
Db 101 ETMFL 105

RESULT 40
C98274
nikP2 protein (AJ250581) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 12-Jul-2004
C:Accession: C98274
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ourullo, B.; Goldmar, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: C98274
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-254 <KUR>
A:Cross-references: UNIPROT:Q8U9P1; GB:AE007870; PIDN:AAK89717.1; PID:gi15159631; GSPDB:G
C:Genetics:
A:Gene: AGR_L_2306
A:Map position: linear chromosome
C:Superfamily: type II thioesterase, NRPS/PKS/S-PAS type; oleoyl-[acyl-carrier-protein] I

Query Match 25.0%; Score 5; DB 2; Length 254;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
|||||
Db 3 RGAPM 7

RESULT 41
H83954
flagellar protein required for flagellar formation flhI [imported] - Bacillus halodurans
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: H83954
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiraoka, Y.; Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: H83954
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-258 <STO>
A:Cross-references: UNIPROT:Q9XA50; GB:AP001515; GB:BA000004; NID:gi10174886; PIDN:BA0061
C:Experimental source: strain C-125
C:Genetics:
A:Gene: flhI

Query Match 25.0%; Score 5; DB 2; Length 258;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETM 18
|||||
Db 244 MLETM 248

RESULT 42
B81971
lacto-N-neotetraose biosynthesis glycosyl transferase NMA0527 [imported] - Neisseria meningitidis
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: B81971
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morelli, M.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: B81971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-268 <PAR>
A:Cross-references: UNIPROT:Q9JW65; GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB8381
A:Experimental source: serogroup A, strain Z2491
C:Genetics:

A:Gene: lgtA2'; NMA0527
C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 25.0%; Score 5; DB 2; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LETMF 19

Db 121 LETMF 125

RESULT 43

C81027
lacto-N-neotetraose biosynthesis glycosyl transferase lgtB NMB1928 [imported] - *Neisseria meningitidis*
C:Species: *Neisseria meningitidis*
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: C81027

R:Tetterlin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; et al.; Gil, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Venter, A.; et al.

A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: C81027

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-275 <TET>

A:Cross-references: UNIPROT:Q51116; GB:AE002541; GB:AE002098; NID:g7227175; PIDN:AAF4225

A:Experimental source: serogroup B, strain MC58

C:Genetics:

C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 25.0%; Score 5; DB 2; Length 275;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LETMF 19

Db 121 LETMF 125

RESULT 44

S70814
glycosyl transferase B (EC 2.4.-.-) - *Neisseria meningitidis*
C:Species: *Neisseria meningitidis*
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004
C:Accession: S70814

R:Jennings, M.P.; Hood, D.W.; Peak, I.R.A.; Virji, M.; Moxon, E.R.

Mol. Microbiol. 18, 729-740, 1995

A:Title: Molecular analysis of a locus for the biosynthesis and phase-variable expression

A:Reference number: S70812; MUID:96414473; PMID:8817494

A:Accession: S70814

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-275 <JEN>

A:Cross-references: UNIPROT:Q51116; EMBL:U25839; NID:g973183; PIDN:AAC44085.1; PID:g9731

C:Genetics:

A:Gene: lgtB

C:Superfamily: lipopolysaccharide biosynthesis-associated protein

C:Keywords: glycosyltransferase

Query Match 25.0%; Score 5; DB 2; Length 275;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LETMF 19

Db 121 LETMF 125

RESULT 45

S70815

glycosyl transferase E (EC 2.4.-.-) - *Neisseria meningitidis*

C:Species: *Neisseria meningitidis*

C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004

C:Accession: S70815

R:Jennings, M.P.; Hood, D.W.; Peak, I.R.A.; Virji, M.; Moxon, E.R.

Mol. Microbiol. 18, 729-740, 1995

A:Title: Molecular analysis of a locus for the biosynthesis and phase-variable expression

A:Reference number: S70812; MUID:96414473; PMID:8817494

A:Accession: S70815

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-276 <JEN>

A:Cross-references: UNIPROT:Q51117; EMBL:U25839; NID:g973183; PIDN:AAC44086.1; PID:g97318

C:Genetics:

A:Gene: lgtE

C:Superfamily: lipopolysaccharide biosynthesis-associated protein

C:Keywords: glycosyltransferase

Query Match 25.0%; Score 5; DB 2; Length 276;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LETMF 19

Db 117 LETMF 121

RESULT 46

E90589

endonuclease iv (endodeoxyribonuclease iv) [imported] - *Mycoplasma pulmonis* (strain UAB

C:Species: *Mycoplasma pulmonis*

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C:Accession: E90589

R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;

Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulm*

A:Reference number: A99512; MUID:21267165; PMID:11353084

A:Accession: E90589

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-277 <KUR>

A:Cross-references: UNIPROT:Q98PU8; GB:AL445566; PID:gl4090036; PIDN:CAC13794.1; GSPDB:G

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPV 6210

C:Superfamily: deoxyribonuclease IV (phage T4-induced)

Query Match 25.0%; Score 5; DB 2; Length 277;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 MLETM 18

Db 143 MLETM 147

RESULT 47

T27610

hypothetical protein ZC477.8 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
R;Du, Z.

submitted to the EMBL Data Library, November 1995
A;Description: The sequence of *C. elegans* cosmid ZC477.

A;Reference number: Z20392

A;Accession: T27610

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

A;Residues: 1-278 <DUZ>

A;Cross-references: EMBL:U40802; PIDN:AAA81511.1; CESP:ZC477.8

C;Genetics:

A;Gene: CESP:ZC477.8

A;Introns: 77/1

Query Match 25.0%; Score 5; DB 2; Length 278;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5

Db 112 PRGAP 116

RESULT 48

A81971

lacto-N-neotetraose biosynthesis glycosyl transferase NMA0525 [imported] - *Neisseria meningitidis*

C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004

C;Accession: A81971

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, N.; Holroyd, S.; Jørgensen, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.

A;Reference number: A81775; UID:20222556; PMID:10761919

A;Accession: A81971

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-279 <PAB>

A;Cross-references: UNIPROT:P57033; GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB8381

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: lgtB; NMA0525

C;Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 25.0%; Score 5; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19

Db 121 LETMF 125

RESULT 49

A81027

lacto-N-neotetraose biosynthesis glycosyl transferase LgtE NMB1926 [imported] - *Neisseria meningitidis*

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C;Accession: A81027

R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizzia, M. Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Vitti, A.; Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.

A;Reference number: A81000; UID:20175755; PMID:10710307

A;Accession: A81027

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-280 <TET>

A;Cross-references: UNIPROT:Q51117; GB:AE002541; GB:AE002098; NID:g7227175; PIDN:AAF4225

A;Experimental source: serogroup B, strain MC58

C;Genetics:

A;Gene: NMB1926

C;Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 25.0%; Score 5; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19

Db 121 LETMF 125

RESULT 50

E70745

hypothetical protein Rv0498 - *Mycobacterium tuberculosis* (strain H37RV)

C;Species: *Mycobacterium tuberculosis*

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C;Accession: E70745

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Reltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome

A;Reference number: A70500; UID:98295987; PMID:9634230

A;Accession: E70745

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-280 <COL>

A;Cross-references: UNIPROT:Q11163; GB:Z77162; GB:AL123456; NID:g3261606; PIDN:CAB00924

A;Experimental source: strain H37RV

C;Genetics:

A;Gene: Rv0498

C;Superfamily: Streptomyces coelicolor hypothetical protein SCE7.14c

Query Match 25.0%; Score 5; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCQML 15

Db 235 VCQML 239

RESULT 51

T22501

hypothetical protein F52D10.4 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T22501

R;Kershaw, J.

submitted to the EMBL Data Library, November 1995

A;Reference number: Z19571

A;Accession: T22501

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-284 <WIL>

A;Cross-references: UNIPROT:Q20656; EMBL:Z66564; PIDN:CAA91475.1; GSPDB:GN00028; CESP:F52D10

C;Genetics:

A;Gene: CESP:F52D10.4

A;Map position: X

A;Introns: 60/3; 104/3; 143/1; 180/3; 255/3

Query Match 25.0%; Score 5; DB 2; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17

Db 180 QMLET 184

RESULT 52

AB1529
transcription regulator (repressor) homolog lin0770 [imported] - *Listeria innocua* (strain C:Species: *Listeria innocua*)
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB1529
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H. D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournan, A.; Mak, C.; Schluter, J.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative Genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1529
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-288 <GLA>
A:Cross-references: UNIPROT:Q92DP3; GB:AL592022; PIDN:CAC96002.1; PID:gl6413221; GSPDB:G
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lin0770
C:Superfamily: glucose kinase; glucose kinase homology

Query Match 25.0%; Score 5; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
|||||
DB 209 LETMF 213

RESULT 53

C72378
sugar ABC transporter, permease protein - *Thermotoga maritima* (strain MSB8)
C:Species: *Thermotoga maritima*
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: C72378
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: C72378
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-294 <ARN>
A:Cross-references: UNIPROT:Q9WY00; GB:AE001721; GB:AE000512; NID:g4980922; PIDN:AAD3550
A:Genetics:
A:Gene: TW0419
C:Superfamily: inner membrane protein ugpA

Query Match 25.0%; Score 5; DB 2; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMWM 8
|||||
DB 149 APMWM 153

RESULT 54

T24827
hypothetical protein T11B7.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24827
R:Gardner, A.
submitted to the EMBL Data Library, September 1995

A:Reference number: Z19940

A:Accession: T24827

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-296 <WIL>

A:Cross-references: UNIPROT:Q22389; EMBL:Z54237; PIDN:CRA90989.1; GSPDB:GN00022; CESP:T11

A:Experimental source: clone T11B7

C:Genetics:

A:Gene: CESP:T11B7.3

A:Map position: 4

A:Introns: 45/3; 111/1; 270/1

Query Match 25.0%; Score 5; DB 2; Length 296;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5

|||||

DB 250 PRGAP 254

RESULT 55

G72776

hypothetical protein APE0201 - *Aeropyrum pernix* (strain K1)

C:Species: *Aeropyrum pernix*

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C:Accession: G72776

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaha

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic *Crenarchaeon*, *Aeropyr*

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: G72776

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-304 <KAW>

A:Cross-references: UNIPROT:Q9YFP8; DDBJ:AP000058; NID:G5103388; PIDN:BAA79113.1; PID:G10

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0201

Query Match 25.0%; Score 5; DB 2; Length 304;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6

|||||

DB 156 RGAPM 160

RESULT 56

T20906

hypothetical protein F14F7.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T20906

R:McMurray, A.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19345

A:Accession: T20906

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-305 <WIL>

A:Cross-references: UNIPROT:O17805; EMBL:Z81503; PIDN:CAB04111.1; GSPDB:GN00021; CESP:F14

A:Experimental source: clone F14F7

C:Genetics:

A:Gene: CESP:F14F7.1

A:Map position: 3

A:Introns: 27/3; 49/3

Query Match 25.0%; Score 5; DB 2; Length 305;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 |||||
 Db 193 PRGAP 197

RESULT 57
 A:Gene: fabD
 A:Accession: A89896
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-308 <KUP>
 A:Cross-references: UNIPROT:Q99UN8; GB:BA000018; PID:G13701030; PIDN:BAB42325.1; GSPDB:G
 A:Experimental source: strain N315
 C:Genetics:
 C:Superfamily: [acyl-carrier-protein] S-malonyltransferase; [acyl-carrier-protein] S-mal

Query Match 25.0%; Score 5; DB 2; Length 308;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
 |||||
 Db 45 LETMF 49

RESULT 58
 H75049
 hypothetical protein PAB1459 - Pyrococcus abyssi (strain Orsay)
 C:Species: Pyrococcus abyssi
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 A:Accession: H75049
 R:anonymous, Genoscope
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
 A:Reference number: A75001
 A:Accession: H75049
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-308 <KAW>
 A:Cross-references: UNIPROT:Q9UYW9; GB:AJ248287; GB:AL096836; NID:G5458657; PIDN:CAB5029
 A:Experimental source: strain Orsay
 C:Genetics:
 C:Superfamily: Pyrococcus abyssi hypothetical protein PAB1459

Query Match 25.0%; Score 5; DB 2; Length 308;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
 |||||
 Db 181 RGAPM 185

RESULT 59
 G71110
 hypothetical protein PH0654 - Pyrococcus horikoshii
 C:Species: Pyrococcus horikoshii
 C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
 A:Accession: G71110
 R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
 DNA Res. 5, 55-76, 1998
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
 A:Reference number: A71000; MUID:98344137; PMID:9679194
 A:Accession: G71110
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-308 <KAW>
 A:Cross-references: UNIPROT:O58188; GB:AP000003; NID:G3236130; PIDN:BAA29745.1; PID:G325
 A:Experimental source: strain OT3
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C:Genetics:
 A:Gene: PH0654

Query Match 25.0%; Score 5; DB 2; Length 308;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
 |||||
 Db 181 RGAPM 185

RESULT 60
 A24815
 calpain (EC 3.4.22.17) large chain 1 - rabbit (fragments)
 N:Alternate names: calcium-activated neutral proteinase (CANP)
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 24-Jan-1988 #sequence_revision 24-Feb-1994 #text_change 09-Jul-2004
 A:Accession: A24815; A41418
 C:Accession: A24815; A41418
 J:Emori, Y.; Kawasaki, H.; Sugihara, H.; Imajoh, S.; Kawashima, S.; Suzuki, K.
 J. Biol. Chem. 261, 9465-9471, 1986
 A:Title: Isolation and sequence analyses of cDNA clones for the large subunits of two iso
 A:Reference number: A92594; MUID:86250902; PMID:2424911
 A:Accession: A24815
 A:Molecule type: mRNA
 A:Residues: 19-320 <EMO>
 A:Cross-references: UNIPROT:P06815; GB:M13363; NID:G165667; PIDN:AAA31456.1; PID:G165668
 R:Kawasaki, H.; Imajoh, S.; Suzuki, K.
 J. Biochem. 102, 393-400, 1987

A:Title: Separation of peptides on the basis of the difference in positive charge: simul
 A:Reference number: A41418; MUID:88032960; PMID:3667575
 A:Accession: A41418
 A:Molecule type: protein
 A:Residues: 1-18;125-154;313-320 <KAW>
 A:Note: sequence was deduced from composition by homology
 C:Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
 C:Keywords: blocked amino end; calcium binding; cysteine proteinase; duplication; EF han
 F:148-179/Domain: calmodulin repeat homology <EF1>
 F:191-223/Domain: calmodulin repeat homology <EF2>
 F:224-253/Domain: calmodulin repeat homology <EF3>
 F:256-288/Domain: calmodulin repeat homology <EF4>
 F:289-320/Domain: calmodulin repeat homology <EF5>

Query Match 25.0%; Score 5; DB 2; Length 320;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
 |||||
 Db 288 LETMF 292

RESULT 61
 AB2574
 hypothetical protein alr9017 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120esp
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 A:Accession: AB2574
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakaazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB2574
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-325 <KUR>
A;Cross-references: UNIPROT:Q8YJU6; GB:AP003605; PIDN:BAH77503.1; PID:g17134948; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr9017
A;Genome: plasmid

Query Match 25.0%; Score 5; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
|||||
Db 228 QMLET 232

RESULT 62
H82455
IS5 transposase VCA0472 VCA0282 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1
C;Species: *Vibrio cholerae*
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: H82455; D82478
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Esmolaeva, M.D.; Vamathevan, J.; Bais, S.; Qin, H.; Dragoi, I.; Sellers, F.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A;Reference number: A82035; MUID:2046833; PMID:10952301
A;Accession: H82455
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-326 <HEI>
A;Cross-references: UNIPROT:Q9K2I8; GB:AE004379; GB:AE003853; NID:g9657865; PIDN:AAF9637
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
A;Accession: D82478
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-326 <HE2>
A;Cross-references: GB:AE004368; GB:AE003853; NID:g9657677; PIDN:AAF96191.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics: <GEN1>
A;Gene: VCA0472
A;Map position: 2
C;Genetics: <GEN2>
A;Gene: VCA0282
A;Map position: 2
C;Superfamily: transposase IS5

Query Match 25.0%; Score 5; DB 2; Length 326;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
|||||
Db 57 LETMF 61

RESULT 63
T02347
probable lipid transfer protein T8F5.2 - *Arabidopsis thaliana*
C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
C;Accession: T02347
R;Vysockaja, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Liu, S.; Li, J.; Arau
Li, Y.; Palm, C.O.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N.A.; The
submitted to the EMBL Data Library, July 1998
A;Description: *Arabidopsis thaliana* chromosome 1 BAC T8F5 complete sequence.
A;Reference number: Z14666

A;Accession: T02347
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-330 <VIS>
A;Cross-references: UNIPROT:O80796; EMBL:AC004512; NID:g3335331; PID:g3335333; PIDN:AAAC2;
C;Genetics:
A;Gene: ATSP.T8F5.2
A;Map position: 1
A;Introns: 47/1; 83/3; 120/3; 142/2; 169/3; 194/3; 218/2; 246/1; 267/3; 292/3
C;Superfamily: probable lipid transfer protein M30

Query Match 25.0%; Score 5; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
|||||
Db 269 QMLET 273

RESULT 64
A40005
hyoxycamine (6S)-dioxygenase (EC 1.14.11.11) - *henbane*
C;Species: *Hyoscyamus niger* (henbane)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A40005
R;Matsuda, J.; Okabe, S.; Hashimoto, T.; Yamada, Y.
J. Biol. Chem. 266, 9460-9464, 1991
A;Title: Molecular cloning of hyoscyamine 6beta-hydroxylase, a 2-oxoglutarate-dependent (C
A;Reference number: A40005; MUID:91236712; PMID:2033047
A;Accession: A40005
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-344 <NAT>
A;Cross-references: UNIPROT:P24397; GB:M62719; NID:g168267; PIDN:AAA33387.1; PID:g168268
C;Superfamily: 1-aminocyclopropane-1-carboxylate oxidase
C;Keywords: ascorbic acid; iron; metalloprotein; oxidoreductase
F;66,217,274/Binding site: iron (His) #status predicted

Query Match 25.0%; Score 5; DB 1; Length 344;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETM 18
|||||
Db 73 MLETM 77

RESULT 65
D96761
unknown protein [imported] - *Arabidopsis thaliana*
C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D96761
R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96761
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-351 <STO>
A;Cross-references: UNIPROT:Q9FX40; GB:AE005173; NID:gl1120789; PIDN:AG30969.1; GSPDB:GN
C;Genetics:
A;Gene: T9L24.34
A;Map position: 1

```

Query Match      25.0%; Score 5; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
DB 321 QMLET 325

RESULT 66
T29932
hypothetical protein F29B9.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29932
R:Goela, D.; Gattung, S.
submitted to the EMBL Data Library, September 1996
A:Description: The sequence of C. elegans cosmid F29B9.
A:Reference number: Z20710
A:Accession: T29932
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-355 <GOE>
A:Cross-references: UNIPROT:Q9GYI3; EMBL:U70849; PIDN:AA09121.1; GSPDB:GN000022; CESP:F2
A:Experimental source: strain Bristol N2; clone F29B9
C:Genetics:
A:Gene: CESP:F29B9.9
A:Map position: 4
A:Introns: 92/1

Query Match      25.0%; Score 5; DB 2; Length 355;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
DB 175 PRGAP 179

RESULT 67
T24137
hypothetical protein R10E4.2a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Aug-2004
C:Accession: T24137
R:Ainscough, R.
submitted to the EMBL Data Library, August 1995
A:Reference number: Z19843
A:Accession: T24137
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-357 <WIL>
A:Cross-references: UNIPROT:Q21900; EMBL:Z50874; PIDN:CAA90772.1; GSPDB:GN000021; CESP:R1
A:Experimental source: clone R10E4
C:Genetics:
A:Gene: CESP:R10E4.2a
A:Map position: 3
A:Introns: 42/3; 80/3; 249/3; 349/3
C:Superfamily: ribonucleoprotein repeat homology

Query Match      25.0%; Score 5; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
DB 176 QMLET 180

RESULT 68
T08903
hypothetical protein T32A16.10 - Arabidopsis thaliana

```

```

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T08903
R:Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May
submitted to the Protein Sequence Database, May 1999
A:Reference number: Z16518
A:Accession: T08903
A:Molecule type: DNA
A:Residues: 1-364 <BEV>
A:Cross-references: UNIPROT:Q9T099; EMBL:AL078468; GSPDB:GN000062; ATSP:T32A16.10
A:Experimental source: cultivar Columbia; BAC clone T32A16
C:Genetics:
A:Gene: ATSP:T32A16.10
A:Map position: 4
A:Introns: 153/1; 325/1

Query Match      25.0%; Score 5; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
DB 60 LETMF 64

RESULT 69
T40115
uv excision repair protein rad23 homolog - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Aug-2004
C:Accession: T40115; T51298
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Taylor, K.; Harris, D.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z21906
A:Accession: T40115
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-368 <WOO>
A:Cross-references: UNIPROT:Q74803; EMBL:AL031788; PIDN:CAA21170.1; GSPDB:GN000067; SPDB:
A:Experimental source: strain 972h-; cosmid C2D10
R:Zhao, Y.; Elder, R.T.
submitted to the EMBL Data Library, July 1999
A:Description: A fission yeast orthologue (rhp23) of the human nucleotide excision repair
A:Reference number: Z25362
A:Accession: T51298
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-368 <ZHA>
A:Cross-references: EMBL:AF174293; PIDN:AA051975.1
A:Experimental source: strain SP223
C:Genetics:
A:Gene: rhp23; SPDB:SPBC2D10.12
A:Map position: 2
A:Introns: 23/3; 48/1; 328/3; 351/3
C:Superfamily: ubiquitin homology

Query Match      25.0%; Score 5; DB 2; Length 368;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
DB 268 QMLET 272

RESULT 70
D82264
probable exopolysaccharide biosynthesis protein EpsF VC0920 [imported] - Vibrio cholerae
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: D82264
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.

```

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A;Reference number: A82035; MUID:20406833; PMID:10952301
 A;Accession: D82264
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-382 <HEI>
 A;Cross-references: UNIPROT:Q9KTI2; GB:AE004175; GB:AE003852; NID:g9655366; PIDN:AAF9408
 A;Experimental source: serogroup O1; strain N18961; biotype El Tor
 C;Genetics:
 A;Gene: VC0920
 A;Map position: 1

Query Match 25.0%; Score 5; DB 2; Length 382;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCQML 15
 |||||
 Db 26 VCQML 30

RESULT 71
 BWHXD
 bexD protein - Haemophilus influenzae
 C;Species: Haemophilus influenzae
 C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
 C;Accession: S12232
 R;Kroll, J.S.; Loynds, B.; Brophy, L.N.; Moxon, E.R.
 Mol. Microbiol. 4, 1853-1862, 1990
 A;Title: The bex locus in encapsulated Haemophilus influenzae: a chromosomal region involved in TDP-daunosamine formation.
 A;Reference number: S12232; MUID:91186821; PMID:2082145
 A;Accession: S12232
 A;Molecule type: DNA
 A;Residues: 1-394 <KRO>
 A;Cross-references: UNIPROT:P22236; EMBL:X54987; NID:g45295; PIDN:CAA38730.1; PID:g45296
 C;Genetics:
 A;Gene: bexD
 C;Superfamily: bexD protein

Query Match 25.0%; Score 5; DB 1; Length 394;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETMFL 20
 |||||
 Db 344 ETMFL 348

RESULT 72
 S39965
 hypothetical protein - Streptomyces griseus
 C;Species: Streptomyces griseus
 C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
 C;Accession: S39965
 R;Kruegel, H.; Schumann, G.; Haenel, F.; Fiedler, G.
 Mol. Gen. Genet. 241, 193-202, 1993
 A;Title: Nucleotide sequence analysis of five putative Streptomyces griseus genes, one of which is involved in TDP-daunosamine formation.
 A;Reference number: S39963; MUID:94049680; PMID:8232204
 A;Accession: S39965
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-406 <KRU>
 A;Cross-references: UNIPROT:Q54197; EMBL:X73149; NID:g407882; PIDN:CAA51670.1; PID:g40788
 C;Superfamily: acyl-CoA dehydrogenase

Query Match 25.0%; Score 5; DB 2; Length 406;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5

Db 242 PRGAP 246

|||||

RESULT 73

T00693

hypothetical protein At2g44130 [imported] - Arabidopsis thaliana

N;Alternate names: hypothetical protein F6E13.26

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004

C;Accession: T00693; G84874

R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, E.
 submitted to the EMBL Data Library, June 1998

A;Description: Arabidopsis thaliana chromosome II BAC F6E13 genomic sequence.

A;Reference number: Z14180

A;Accession: T00693

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-407 <ROU>

A;Cross-references: UNIPROT:O80582; EMBL:AC004005; NID:g3212846; PID:g3212866

A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N.
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: G84874

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-407 <STO>

A;Cross-references: GB:AE002093; NID:g3212866; PIDN:AAC23417.1; GSPDB:GN00139

C;Genetics:

A;Gene: F6E13.26; At2g44130

A;Map position: 2

Query Match

25.0%; Score 5; DB 2; Length 407;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6

|||||

Db 179 RGAPM 183

RESULT 74

T24138

hypothetical protein R10E4.2b - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Aug-2004

C;Accession: T24138

R;Ainscough, R.

submitted to the EMBL Data Library, August 1995

A;Reference number: Z19843

A;Accession: T24138

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-409 <WIL>

A;Cross-references: UNIPROT:Q21900; EMBL:Z50874; PIDN:CAA90773.1; GSPDB:GN00021; CESP:R10
 A;Experimental source: clone R10E4

C;Genetics:

A;Gene: CESP:R10E4.2b

A;Map position: 3

A;Introns: 42/3; 106/3; 301/3; 401/3

C;Superfamily: ribonucleoprotein repeat homology

Query Match

25.0%; Score 5; DB 2; Length 409;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17

|||||

Db 202 QMLET 206

RESULT 75

D87492
Hypothetical protein CC1961 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: D87492
R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: D87492
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-420 <STO>
A;Cross-references: UNIPROT:O87708; GB:AE005673; NID:gl3423422; PIDN:AAK23936.1; GSPDB:G
C;Genetics:
A;Gene: CC1961
C;Superfamily: heat shock protein hslU; FtsH/SEC18/CDC48-type ATP-binding domain homolog

Query Match 25.0%; Score 5; DB 2; Length 420;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
|||
Db 378 LETMF 382

Search completed: October 26, 2004, 07:23:36
Job time : 20.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:48:06 ; Search time 91.5 Seconds
(without alignments)
125.765 Million cell updates/sec

Title: US-10-066-965A-3

Perfect score: 20

Sequence: 1 PRGAPWVRWVCOMLETWPL 20

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1825181 seqs, 575374646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Uniprot 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	30.0	82	Q8VJN1	Q8vjn1 mycobacteri
2	6	30.0	189	Q6VAB5	Q6vab5 trichosurus
3	6	30.0	189	AAQ24229	AAQ24229 trichosur
4	6	30.0	192	DEF_TETH	DEF_TETH thermus t
5	6	30.0	192	Q72H33	Q72h33 thermus t
6	6	30.0	192	AAS82004	AAS82004 thermus t
7	6	30.0	232	DEF_DEIRA	DEF_DEIRA deincococ
8	6	30.0	253	Q8LNA8	Q8lna8 oryza sativ
9	6	30.0	254	Q83VQ5	Q83vq5 methylobaci
10	6	30.0	341	Q7NNH5	Q7nnh5 chromobacte
11	6	30.0	352	POSI_MOUSE	P20263 mus musculu
12	6	30.0	352	AAH68268	AAH68268 mus muscu
13	6	30.0	384	Q9C280	Q9c280 neurospora
14	6	30.0	411	Q87IP7	Q87ip7 vibrio para
15	6	30.0	438	Q8IV69	Q8iv69 homo sapien
16	6	30.0	478	Q6FLU1	Q6flu1 mesoplasma
17	6	30.0	478	Q6MTY4	Q6mt4 mycoplasma
18	6	30.0	478	CAE76902	CAE76902 mycoplasm
19	6	30.0	758	Q7KJB9	Q7kjb9 drosophila
20	6	30.0	1037	Q7YV13	Q7yv13 trypanosoma
21	6	30.0	1345	Q8A180	Q8a180 bacteroides
22	6	30.0	1437	1 VGLM_BUNGE	Pl2430 bunyavirus
23	6	30.0	1703	Q9V9A9	Q9v9a9 drosophila
24	5	25.0	23	Q83SX5	Q83sx5 salmonella
25	5	25.0	28	Q80KE5	Q80ke5 rabies viru
26	5	25.0	28	Q80KE7	Q80ke7 rabies viru
27	5	25.0	28	Q80KE9	Q80ke9 rabies viru
28	5	25.0	28	Q80KF0	Q80kf0 rabies viru
29	5	25.0	28	Q80KF2	Q80kf2 rabies viru
30	5	25.0	28	Q80KF3	Q80kf3 rabies viru
31	5	25.0	28	Q80KF4	Q80kf4 rabies viru

32	5	25.0	28	Q80KF5	Q80kf5 rabies viru
33	5	25.0	28	Q80KF6	Q80kf6 rabies viru
34	5	25.0	28	Q80KF8	Q80kf8 rabies viru
35	5	25.0	28	Q80KF9	Q80kf9 rabies viru
36	5	25.0	28	Q80KG1	Q80kg1 rabies viru
37	5	25.0	28	Q80KG2	Q80kg2 rabies viru
38	5	25.0	28	Q80KG3	Q80kg3 rabies viru
39	5	25.0	28	Q80KG4	Q80kg4 rabies viru
40	5	25.0	28	Q80KG7	Q80kg7 rabies viru
41	5	25.0	28	Q80KG8	Q80kg8 rabies viru
42	5	25.0	28	Q80KG9	Q80kg9 rabies viru
43	5	25.0	28	Q80KH0	Q80kh0 rabies viru
44	5	25.0	28	Q80KH1	Q80kh1 rabies viru
45	5	25.0	28	Q80KJ9	Q80kj9 rabies viru
46	5	25.0	28	Q80KK1	Q80kk1 rabies viru
47	5	25.0	28	Q80KK2	Q80kk2 rabies viru
48	5	25.0	28	Q80KK4	Q80kk4 rabies viru
49	5	25.0	28	Q80KK5	Q80kk5 rabies viru
50	5	25.0	28	Q80KK6	Q80kk6 rabies viru
51	5	25.0	28	Q80KK7	Q80kk7 rabies viru
52	5	25.0	28	Q80KK8	Q80kk8 rabies viru
53	5	25.0	28	Q80KK9	Q80kk9 rabies viru
54	5	25.0	28	Q80KL0	Q80kl0 rabies viru
55	5	25.0	28	Q80KL1	Q80kl1 rabies viru
56	5	25.0	28	Q80KL2	Q80kl2 rabies viru
57	5	25.0	28	Q80KL3	Q80kl3 rabies viru
58	5	25.0	28	Q80KL5	Q80kl5 rabies viru
59	5	25.0	28	Q9IWC4	Q9iwc4 rabies viru
60	5	25.0	28	Q9IWC5	Q9iwc5 rabies viru
61	5	25.0	28	Q9QD10	Q9qd10 rabies viru
62	5	25.0	28	Q9QD11	Q9qd11 rabies viru
63	5	25.0	28	Q9QD12	Q9qd12 rabies viru
64	5	25.0	28	Q9QD13	Q9qd13 rabies viru
65	5	25.0	28	Q9QD14	Q9qd14 rabies viru
66	5	25.0	28	Q9QD15	Q9qd15 rabies viru
67	5	25.0	28	Q9QD16	Q9qd16 rabies viru
68	5	25.0	28	Q9QD17	Q9qd17 rabies viru
69	5	25.0	28	Q9QD18	Q9qd18 rabies viru
70	5	25.0	28	Q9QD19	Q9qd19 rabies viru
71	5	25.0	28	Q9QD22	Q9qd22 rabies viru
72	5	25.0	28	Q9QD23	Q9qd23 rabies viru
73	5	25.0	28	Q9QD24	Q9qd24 rabies viru
74	5	25.0	28	Q9QD25	Q9qd25 rabies viru
75	5	25.0	28	Q9QD26	Q9qd26 rabies viru
76	5	25.0	28	Q9QD27	Q9qd27 rabies viru
77	5	25.0	28	Q9EIT2	Q9elt2 rabies viru
78	5	25.0	28	Q9EIT3	Q9elt3 rabies viru
79	5	25.0	28	Q9EIT4	Q9elt4 rabies viru
80	5	25.0	28	Q9EIT5	Q9elt5 rabies viru
81	5	25.0	28	Q9EIT6	Q9elt6 rabies viru
82	5	25.0	28	Q9EIT7	Q9elt7 rabies viru
83	5	25.0	28	Q9E2R9	Q9e2r9 rabies viru
84	5	25.0	28	Q9E2S0	Q9e2s0 rabies viru
85	5	25.0	28	Q9E2S1	Q9e2s1 rabies viru
86	5	25.0	28	Q9E2S2	Q9e2s2 rabies viru
87	5	25.0	28	Q9E2S3	Q9e2s3 rabies viru
88	5	25.0	28	Q6DQ19	Q6dql9 rabies viru
89	5	25.0	28	Q6DQ10	Q6dql0 rabies viru
90	5	25.0	28	Q6DQ11	Q6dql1 rabies viru
91	5	25.0	28	Q6DQ12	Q6dql2 rabies viru
92	5	25.0	28	Q6DQ13	Q6dql3 rabies viru
93	5	25.0	28	Q6DQ14	Q6dql4 rabies viru
94	5	25.0	28	Q6DQ15	Q6dql5 rabies viru
95	5	25.0	28	Q6DQ16	Q6dql6 rabies viru
96	5	25.0	28	Q6DQ17	Q6dql7 rabies viru
97	5	25.0	28	Q6DQ18	Q6dql8 rabies viru
98	5	25.0	28	Q6DQ19	Q6dql9 rabies viru
99	5	25.0	28	Q6DQ20	Q6dql0 rabies viru
100	5	25.0	28	Q6DQ21	Q6dql1 rabies viru

ALIGNMENTS

```

RESULT 1
Q8VJN1
ID Q8VJN1 PRELIMINARY; PRT; 82 AA.
AC Q8VJN1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein MT2341.
GN OrderedLocNames=MT2341;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL; AB007077; AAK46625.1; -.
DR TIGR; MT2341; -.
KW Hypothetical protein.
SQ SEQUENCE 82 AA; 8995 MW; C46009E4E66DC09E CRC64;

Query Match 30.0%; Score 6; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPMW 7
Db |||||
6 RGAPMW 11

RESULT 2
Q6VAB5
ID Q6VAB5 PRELIMINARY; PRT; 189 AA.
AC Q6VAB5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE POU domain class 5 transcription factor 1-like protein
DE (Fragment).
GN Name=POU5F1;
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
OX NCBI_TaxID=9337;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21094456; PubMed=11170265;
RA Frankenberg S., Tisdall D., Selwood L.;
RT "Identification of a homologue of POU5F1 (OCT3/4) in a marsupial, the
RT brushtail possum.";
RL Mol. Reprod. Dev. 58:255-261(2001).
RN [2]
RN SEQUENCE FROM N.A.
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
OX NCBI_TaxID=9337;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21094456; PubMed=11170265;
RA Frankenberg S., Tisdall D., Selwood L.;
RT "Identification of a homologue of POU5F1 (OCT3/4) in a marsupial, the
RT brushtail possum.";
RL Mol. Reprod. Dev. 58:255-261(2001).
RN [2]
RN SEQUENCE FROM N.A.
RA Frankenberg S., Tisdall D., Selwood L.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AY345973; AAQ24229.1; -.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; Homeobox.
DR InterPro; IPR010982; Lambda_like_DNA.
DR InterPro; IPR000327; POU.
DR InterPro; IPR007103; POU_homeo.

```

```

DR Pfam; PF00046; Homeobox; 1.
DR Pfam; PF00157; Pou; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00028; POU DOMAIN.
DR ProDom; PD000010; Homeobox; 1.
DR ProDom; PD000583; POU; 1.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00352; POU; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR PROSITE; PS00465; POU_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
SQ SEQUENCE 189 AA; 21027 MW; 119434E565647C6B CRC64;

Query Match 30.0%; Score 6; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMFL 20
Db |||||
76 LETMFL 81

RESULT 3
AAQ24229
ID AAQ24229 PRELIMINARY; PRT; 189 AA.
AC AAQ24229;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE POU domain class 5 transcription factor 1-like protein
DE (Fragment).
GN POU5F1.
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
OX NCBI_TaxID=9337;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21094456; PubMed=11170265;
RA Frankenberg S., Tisdall D., Selwood L.;
RT "Identification of a homologue of POU5F1 (OCT3/4) in a marsupial, the
RT brushtail possum.";
RL Mol. Reprod. Dev. 58:255-261(2001).
RN [2]
RN SEQUENCE FROM N.A.
RA Frankenberg S., Tisdall D., Selwood L.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY345973; AAQ24229.1; -.
FT NON_TER 1
SQ SEQUENCE 189 AA; 21027 MW; 119434E565647C6B CRC64;

Query Match 30.0%; Score 6; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMFL 20
Db |||||
76 LETMFL 81

RESULT 4
DEF_TETH
ID DEF_TETH STANDARD; PRT; 192 AA.
AC P43522;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN Name=def;
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;

```



```
OC Thermus
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VK1;
RC MEDLINE=95050326; PubMed=7961514;
RA Meinel T., Blanquet S.;
RT "Characterization of the Thermus thermophilus locus encoding peptide
  deformylase and methionyl-tRNA(Met) formyltransferase.";
RL J. Bacteriol. 176:7387-7390 (1994).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=97272005; PubMed=9126850;
RA Meinel T., Lazemec C., Viljoing S., Blanquet S.;
RT "Structure-function relationships within the peptide deformylase
  family. Evidence for a conserved architecture of the active site
  involving three conserved motifs and a metal ion.";
RL J. Mol. Biol. 267:749-761 (1997).
CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
  newly synthesized proteins. Requires at least a dipeptide for an
  efficient rate of reaction. N-terminal L-methionine is a
  prerequisite for activity but the enzyme has broad specificity at
  other positions (By similarity).
CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
  methionyl peptide.
CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@lsb-sib.ch).
-----
CC EMBL; X79087; CAA55695.1; -.
CC HSSP; P96113; 1LME.
CC HAMAP; MF_00163; -.
CC InterPro; IPR000181; Fmet_deformylase.
CC Pfam; PF01327; Pep_deformylase; 1.
CC PRINTS; PR01576; PDEFORMLYASE.
CC ProDom; PD003844; Pep_deformylase; 1.
CC TIGRFAMs; TIGR00079; Pept_deformyl; 1.
KW Hydrolase; Iron; Protein Biosynthesis.
FT ACT_SITE 146 146 By similarity.
FT METAL 102 102 Iron (By similarity).
FT METAL 145 145 Iron (By similarity).
FT METAL 149 149 Iron (By similarity).
SQ SEQUENCE 192 AA; 22092 MW; 665945183A251361 CRC64;

Query Match 30.0%; Score 6; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETFM 19
DB 33 MLETFM 38

RESULT 5
Q72H33 PRELIMINARY; PRT; 192 AA.
AC Q72H33;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Polypeptide deformylase, fms protein-like protein (EC 3.5.1.31).
GN OrderedLocusNames=TTC1662;
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=262724;

Query Match 30.0%; Score 6; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETFM 19
DB 33 MLETFM 38

RESULT 5
Q72H33 PRELIMINARY; PRT; 192 AA.
AC Q72H33;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Polypeptide deformylase, fms protein-like protein (EC 3.5.1.31).
GN OrderedLocusNames=TTC1662;
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=262724;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15064768;
RA Henne A., Brueggemann H., Raasch C., Wierzer A., Hartsch T.,
  Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
  Jacobbi C., Starkuviene V., Schlenczek S., Dencker S., Huber R.,
  Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;
RT "The genome sequence of the extreme thermophile Thermus
  thermophilus.";
RL Nat. Biotechnol. 22:547-553 (2004).
CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.
DR EMBL; AE017306; AAS82004.1; -.
DR GO; GO:0008463; F:formylmethionine deformylase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000181; Fmet_deformylase.
DR Pfam; PF01327; Pep_deformylase; 1.
DR PRINTS; PR01576; PDEFORMLYASE.
DR ProDom; PD003844; Fmet_deformylase; 1.
DR TIGRFAMs; TIGR00079; Pept_deformyl; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 192 AA; 22092 MW; 665945183A251361 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETFM 19
DB 33 MLETFM 38

RESULT 6
AAS82004 PRELIMINARY; PRT; 192 AA.
AC AAS82004;
DT 14-APR-2004 (TrEMBLrel. 27, Created)
DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 11-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Polypeptide deformylase, fms protein-like protein (EC 3.5.1.31).
GN TTC1662.
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=262724;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15064768;
RA Henne A., Brueggemann H., Raasch C., Wierzer A., Hartsch T.,
  Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
  Jacobbi C., Starkuviene V., Schlenczek S., Dencker S., Huber R.,
  Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;
RT "The genome sequence of the extreme thermophile Thermus
  thermophilus.";
RL Nat. Biotechnol. 22:547-553 (2004).
DR EMBL; AE017306; AAS82004.1; -.
KW Hydrolase.
SQ SEQUENCE 192 AA; 22092 MW; 665945183A251361 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETFM 19
DB 33 MLETFM 38

RESULT 7
DEF DEIRA
ID DEF DEIRA STANDARD; PRT; 232 AA.
AC Q9RRQX;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
```

01-OCT-2004 (Rel. 45, Last annotation update)
 DE Peptide deformylase (EC 3.5.1.18) (pDF) (Polypeptide deformylase).
 GN Name=def; OrderedLocusNames=DR2434;
 OS Deinococcus radiodurans
 OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 radiodurans R1.";
 RL Science 286:1571-1577(1999).
 CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
 CC newly synthesized proteins. Requires at least a dipeptide for an
 CC efficient rate of reaction. N-terminal L-methionine is a
 CC prerequisite for activity but the enzyme has broad specificity at
 CC other positions (By similarity).
 CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
 CC methionyl peptide.
 CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
 CC -!- SIMILARITY: Belongs to the polypeptide deformylase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; AE002073; AAF11975.1; -.
 DR PIR; H75274; H75274.
 DR HSP; Q917A8; 1NSN.
 DR TIGR; DR2434; -.
 DR HAMAP; MF_00163; -; 1.
 DR InterPro; IPR000181; Fmet_deformylase.
 DR Pfam; PF01327; Pep_deformylase; 1.
 DR PRINTS; PR01576; PDBFORMYLASE.
 DR ProDom; PD003844; Pep_deformylase; 1.
 DR TIGRFAMs; TIGR00079; pep_deformyl; 1.
 DR Complete proteome; Hydrolase; Iron; Protein biosynthesis.
 FT ACT_SITE 179 179 By similarity.
 FT METAL 135 135 Iron (By similarity).
 FT METAL 178 178 Iron (By similarity).
 FT METAL 182 182 Iron (By similarity).
 SQ SEQUENCE 232 AA; 26369 MW; 94A31606AB872ED8 CRC64;
 Query Match 30.0%; Score 6; DB 1; Length 232;
 Best Local Similarity 100.0%; Pred. No. 63;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 14 MLETFM 19
 DB 62 MLETFM 67
 RESULT 8
 Q8LNA8
 ID Q8LNA8 PRELIMINARY; PRT; 253 AA.
 AC Q8LNA8
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein OSUNBa0047G15.19.

GN ORFNames=OSUNBa0047G15.19;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Embryophyta; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McComb W.R., Spiegel L., de la Bastide M., Nascimento L.,
 RA Preston R., Ferraro K., Kuit K., Zutavern T., Ballia V., Bell M.,
 RA Baker J., Miller B., Katzenberger F., Muller S., King L., Sullivan P.,
 RA Yang C., Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA The Rice Chromosome 10 Sequencing Consortium;
 RT "In-depth view of structure, activity, and evolution of rice
 RT chromosome 10.";
 RL Science 300:1566-1569(2003).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Buell C.R., Wang R.A., McComb W.R., Messing J., Yuan Q.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC083944; XAM74301.1; -.
 DR EMBL; AE017064; AAP52436.1; -.
 DR Gramene; Q8LNA8; -.
 KW Hypothetical protein.
 SQ SEQUENCE 253 AA; 27932 MW; 475202DE0DECB5D CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 253;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GAPMMW 8
 DB 18 GAPMMW 23
 RESULT 9
 Q83VQ5
 ID Q83VQ5 PRELIMINARY; PRT; 254 AA.
 AC Q83VQ5
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Epsp.
 GN Name=epsP;
 OS Methylobacillus sp. 12S.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Methylophilales;
 OC Methylophilaceae; Methylobacillus.
 OX NCBI_TaxID=94001;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=12S;
 RX MEDLINE=22511518; PubMed=12624205;
 RA Yoshida T., Ayabe Y., Yasunaga M., Usami Y., Habe H., Nojiri H.,
 RA Omori T.;
 RT "Genes involved in the synthesis of the exopolysaccharide methanolan
 RT by the obligate methylophilic Methylobacillus sp. strain 12S.";
 RL Microbiology 149:431-444(2003).
 DR EMBL; AB062506; BACS5144.1; -.
 DR GO; GO:0009058; P:biosynthesis; IPA.
 DR InterPro; IPR004629; WecB_TaGA_CpsF.
 DR Pfam; PF03808; Glyco_tran_WecB; 1.
 DR TIGRFAMs; TIGR00696; _wecB_TaGA_CpsF; 1.
 SQ SEQUENCE 254 AA; 28235 MW; AEC67E24A8BE94C3 CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 254;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 APMMWR 9
 |||||

Query Match 30.0%; Score 6; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 6: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

DR EMBL; S58426; AAB19896.1; -.
DR EMBL; S58422; AAB19896.1; JOINED.
DR EMBL; S58423; AAB19896.1; JOINED.
DR EMBL; S58424; AAB19896.1; JOINED.
DR EMBL; S58425; AAB19896.1; JOINED.
DR PIR; A34672; A34672.
DR PIR; S17313; S17313.
DR PDB; 1OCF; NMR; @=216-282.
DR TRANSFAC; T00651; -.
DR MGD; MGI:101893; Pou5f1.
DR GO; GO:0005634; C:nucleus; IC.
DR GO; GO:0003700; F:transcription factor activity; IDA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IDA.
DR GO; GO:0006366; P:transcription from Pol II promoter; IDA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; Homeodomain like.
DR InterPro; IPR010982; Lambda_like_DNA.
DR InterPro; IPR000327; POU.
DR InterPro; IPR007103; POU homeo.
DR Pfam; PF00046; Homeobox; 1.
DR Pfam; PF00157; Pou; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00028; POU DOMAIN.
DR ProDom; PD000010; Homeobox; 1.
DR ProDom; PD000583; POU domain; 1.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00352; POU; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR PROSITE; PS00035; POU 1; 1.
DR PROSITE; PS00465; POU 2; 1.
KW 3D-structure; DNA-binding; Homeobox; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 135 205 POU.
FT DNA_BIND 223 282 Homeobox.
FT CONFLICT 1 28 Missing (in Ref. 2).
FT CONFLICT 29 29 V -> M (in Ref. 2).
FT CONFLICT 31 31 P -> S (in Ref. 4).
FT HELIX 232 242
FT HELIX 250 260
FT TURN 261 261
FT HELIX 264 275
SQ SEQUENCE 352 AA; 38216 MW; 757E41DF52286714 CRC64;

Query Match 30.0%; Score 6; DB 1; Length 352;
Best Local Similarity 100.0%; Pred.No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMFL 20
DB 238 LETMFL 243

RESULT 12
AAH68268
ID AAH68268 PRELIMINARY; PRT; 352 AA.
AC AAH68268;
DT 14-APR-2004 (TrEMBLrel. 27, Created)
DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

Query Match 30.0%; Score 6; DB 2; Length 384;
Best Local Similarity 100.0%; Pred.No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 COMLET 17
DB 185 COMLET 190

RESULT 14
Q87IP7 PRELIMINARY; PRT; 411 AA.
ID Q87IP7

```

```

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC068268; AAB68268.1; -.
KW Hypothetical protein.
SQ SEQUENCE 352 AA; 38216 MW; 757E41DF52286714 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 352;
Best Local Similarity 100.0%; Pred.No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMFL 20
DB 238 LETMFL 243

RESULT 13
Q9C280
ID Q9C280 PRELIMINARY; PRT; 384 AA.
AC Q9C280;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein B13A5.030.
GN Name=B13A5.030;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AL513465; CAC28792.1; -.
DR InterPro; IPR010987; GST_C_like.
KW Hypothetical protein.
SQ SEQUENCE 384 AA; 42398 MW; BB94D063B3D0003B CRC64;

Query Match 30.0%; Score 6; DB 2; Length 384;
Best Local Similarity 100.0%; Pred.No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 COMLET 17
DB 185 COMLET 190

RESULT 14
Q87IP7 PRELIMINARY; PRT; 411 AA.
ID Q87IP7

```

AC Q87IP7;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein VPA0559;
 GN OrderedLocusNames=VPA0559;
 OS Vibrio parahaemolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=670;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RIMD 2210633 / Serotype O3:k6;
 RX MEDLINE=22508454; PubMed=12620739;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
 RT "Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism
 RT distinct from that of *V. cholerae*.";
 RL Lancet 361:743-749(2003).
 DR EMBL; AF005085; BAC61902.1; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR003838; DUF214.
 DR Pfam; PF02687; FtsX; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 411 AA; 45224 MW; 99CB9E2E7E7AD331 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 411;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMFL 20
 |||||
 Db 320 LETMFL 325

RESULT 15
 Q8IV69 PRELIMINARY; PRT; 438 AA.
 AC Q8IV69;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE SLCO1A2 protein (Fragment).
 GN Name=SLCO1A2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.D., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzyzanski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC042452; AAH42452.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR004157; OATP_Cterm.
 DR InterPro; IPR004156; OATP_Nterm.
 DR Pfam; PF03137; OATP; 1.
 DR TIGRFAMS; TIGR00805; oat; 1.
 DR PROSITE; PS50850; MFS; 1.
 FT NON_TER 1
 SQ SEQUENCE 438 AA; 49086 MW; 8367CC253A77D433 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 438;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMFL 20
 |||||
 Db 28 LETMFL 33

RESULT 16
 Q6FIU1 PRELIMINARY; PRT; 478 AA.
 AC Q6FIU1;
 DT 01-OCT-2004 (TrEMBLrel. 28, Created)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Pyruvate kinase.
 GN ORFNames=Mf1175;
 OS Mesoplasma florum L1.
 OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
 OC Entomoplasmataceae; Mesoplasma.
 OX NCBI_TaxID=265311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L1;
 RA Birren B.W., Stange-Thomann N., Hafez N., DeCaprio D., Fisher S.,
 RA Butler J., Elkins T., Kodira C.D., Major J., Wang S., Nicol R.,
 RA Nusbaum C.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L1;
 RA Knight T., Fournier G.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017263; AAT75532.1; -;
 KW Kinase; Pyruvate.
 SQ SEQUENCE 478 AA; 53085 MW; DEF903F6A09938DD CRC64;

Query Match 30.0%; Score 6; DB 2; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLETM 18
 |||||
 Db 290 QMLETM 295

RESULT 17
 Q6MTY4 PRELIMINARY; PRT; 478 AA.
 AC Q6MTY4;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Pyruvate kinase (EC 2.7.1.40).

GN Name=pyk; OrderedlocusNames=MSC_0261;
 OS Mycoplasma mycoides (subsp. mycoides SC).
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=44101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PG1;
 RX PubMed=14762060;
 RA Westberg J., Persson A., Holmberg A., Goessmann A., Lundeberg J.,
 RA Johansson K.-E., Pettersson B., Uhlen M.;
 RT "The genome sequence of Mycoplasma mycoides subsp. mycoides SC type
 strain PG1T, the causative agent of contagious bovine pleuropneumonia
 (CBPP).";
 RT Genome Res. 14:221-227(2004).
 RL Gene! CATALYTIC ACTIVITY: ATP + pyruvate = ADP + phosphoenolpyruvate.
 CC -! COFACTOR: Requires magnesium and potassium (By similarity).
 CC -! PATHWAY: Glycolysis; final step.
 CC -! SUBUNIT: Homotetramer (By similarity).
 CC -! SIMILARITY: Belongs to the pyruvate kinase family.
 DR EMBL; BX842642; CAE76902.1; -.
 DR GO; GO:0004743; F:pyruvate kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR001697; Pyruvate_kinase.
 DR Pfam; PF00224; PK; 1.
 DR PRINTS; PR01050; PYRUVTKNASE.
 DR ProDom; PD001009; Pyruvate_kinase; 2.
 DR TIGRFAMs; TIGR01064; pyruv_kin; 1.
 DR Complete proteome; Glycolysis; Kinase; Magnesium; Pyruvate;
 KW Transferase.
 SQ SEQUENCE 478 AA; 53817 MW; 1CB570E98A6305C9 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLETM 18
 Db 290 QMLETM 295

RESULT 18
 CAE76902
 ID CAE76902 PRELIMINARY; PRT; 478 AA.
 AC CAE76902;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 13-APR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Pyruvate kinase [EC 2.7.1.40].
 GN PYK OR MSC_0261.
 OS Mycoplasma mycoides (subsp. mycoides SC).
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=44101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PG1;
 RX PubMed=14762060;
 RA Westberg J., Persson A., Holmberg A., Goessmann A., Lundeberg J.,
 RA Johansson K.-E., Pettersson B., Uhlen M.;
 RT "The genome sequence of Mycoplasma mycoides subsp. mycoides SC type
 strain PG1T, the causative agent of contagious bovine pleuropneumonia
 (CBPP).";
 RT Genome Res. 14:221-227(2004).
 RL EMBL; BX842642; CAE76902.1; -.
 DR EMBL; BX842642; CAE76902.1; -.
 KW Kinase; Pyruvate; Transferase.
 SQ SEQUENCE 478 AA; 53817 MW; 1CB570E98A6305C9 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLETM 18
 Db 290 QMLETM 295

RESULT 19
 Q7KJB9
 ID Q7KJB9 PRELIMINARY; PRT; 758 AA.
 AC Q7KJB9;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE L(2)01289 long form (fragment).
 GN Name=1(2)01289;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Manning G., Micklem D.R., Krasnow M.A.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF216976; AAF34750.1; -.
 DR InterPro; IPR006663; Thioresox_dom2.
 FT NON_TER 1
 FT NON_TER 758
 SQ SEQUENCE 758 AA; 88075 MW; CB31F4196C4BA315 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 758;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLETM 18
 Db 535 QMLETM 540

RESULT 20
 Q7YV13
 ID Q7YV13 PRELIMINARY; PRT; 1037 AA.
 AC Q7YV13;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Protein kinase, putative.
 GN ORFNames=Tb927.2.4200;
 OS Trypanosoma brucei.
 OC Eukaryota; Euzoenzoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22789168; PubMed=12907728;
 RA El-Sayed N.M.A., Ghedin E., Song J., MacLeod A., Bringaud F.,
 RA Larkin C., Wanless D., Peterson J., Hou L., Taylor S., Tweedie A.,
 RA Biteau N., Khalak H.G., Lin X., Mason T., Hannick L., Calet E.,
 RA Blandin G., Bartholomeu D., Simpson A.J., Kaul S., Zhao H., Pai G.,
 RA Van Aken S., Utterback T., Haas B., Koo H.L., Umayam L., Suh B.,
 RA Gerrard C., Leech V., Qi R., Zhou S., Schwartz D., Feldblyum T.,
 RA Salzberg S., Tait A., Turner M.R., Ullu E., White O., Melville S.,
 RA Adams M.D., Fraser C.M., Donelson J.E.;
 RT "The sequence and analysis of Trypanosoma brucei chromosome II.";
 RL Nucleic Acids Res. 31:4856-4863(2003).
 CC -! SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AE017169; AAQ15877.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1037 AA; 113326 MW; A87B8C4442745274 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 1037;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCOMLE 16
|||
Db 85 VCOMLE 90

RESULT 21

ID Q8A180 PRELIMINARY; PRT; 1345 AA.
AC Q8A180; 2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Two-component system sensor histidine kinase/response regulator,
DE hybrid ('one-component system').
GN OrderedLocusNames=BT3786;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=1263928;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
CC -1- SIMILARITY: Contains 1 HTH arac/xyIS-type DNA-binding domain.
CC -1- SIMILARITY: Contains 1 histidine kinase domain.
EMBL; AE016942; AA078891.1; -;
DR HSP; P08402; I800.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
InterPro; IPR003594; Arpbind ATPase.
DR InterPro; IPR011006; CheY like.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinase.
DR InterPro; IPR000005; HTHarac.
DR InterPro; IPR011110; Reg_prop.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; Hiska; 1.
DR Pfam; PF00165; HTH_Arac; 2.
DR Pfam; PF07494; Reg_prop; 11.
DR Pfam; PF00072; Response_reg; 1.
DR Pfam; PF07495; Y_Y; 1.
DR PRINTS; PR00032; HTHARAC.
DR ProDom; PD0000039; Response_reg; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; Hiska; 1.
DR SMART; SM00342; HTH_Arac; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS01019; HIS_KIN; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
DR PROSITE; PS01110; RESPONSE_REGULATORY; 1.
DR PROSITE; PS01110; RESPONSE_REGULATORY; 1.
KW Complete proteome; DNA-binding; Kinase; Phosphorylation;
KW Sensory transduction; Transcription regulation; transferase.
SQ SEQUENCE 1345 AA; 155269 MW; 7247BDDA9C7FADAB CRC64;

Query Match 30.0%; Score 6; DB 2; Length 1345;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCOMLE 16
|||
Db 1202 VCOMLE 1207

RESULT 22

ID VGLM_BUNGE STANDARD; PRT; 1437 AA.
AC P12430;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE M polyprotein precursor [Contains: Glycoprotein G2; Nonstructural
DE protein NS-M; Glycoprotein G1].
GN Name=M;
OS Bunyavirus germiston.
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Orthobunyavirus.
OX NCBI_TaxID=11574;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89021381; PubMed=3176688;
RA Pardigon N., Vialat P., Gerbaud S., Girard M., Bouloy M.;
RT "Nucleotide sequence of the M segment of Germiston virus: comparison
RT of the M gene product of several bunyaviruses.";
RL Virus Res. 11:73-85(1988).
CC -1- PTM: Specific enzymatic cleavages yield mature proteins including
CC nonstructural protein NS-M, glycoprotein G1, and glycoprotein G2.
CC -1- SIMILARITY: Belongs to the orthobunyaviruses M polyprotein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
EMBL; M21951; AAA42778.1; -;
DR PIR; S07430; S07430.
DR InterPro; IPR005167; Bunya_G1.
DR InterPro; IPR005168; Bunya_G2.
DR Pfam; PF03557; Bunya_G1; 1.
DR Pfam; PF03563; Bunya_G2; 1.
DR Glycoprotein; Nonstructural protein; Polyprotein; Signal;
KW Transmembrane.
FT SIGNAL 1 21 M polyprotein.
FT CHAIN 22 1437 Glycoprotein G2.
FT CHAIN 22 306 Nonstructural protein NS-M.
FT CHAIN 307 481 Glycoprotein G1.
FT CHAIN 482 1437 Potential.
FT TRANSMEM 208 228 Potential.
FT TRANSMEM 316 336 Potential.
FT TRANSMEM 368 388 Potential.
FT TRANSMEM 1395 1415 Potential.
FT CARBOHYD 65 65 N-linked (GLCNAC. . .) (Potential).
FT CARBOHYD 88 88 N-linked (GLCNAC. . .) (Potential).
FT CARBOHYD 252 252 N-linked (GLCNAC. . .) (Potential).
FT CARBOHYD 474 474 N-linked (GLCNAC. . .) (Potential).
FT CARBOHYD 627 627 N-linked (GLCNAC. . .) (Potential).
FT CARBOHYD 1173 1173 N-linked (GLCNAC. . .) (Potential).
SQ SEQUENCE 1437 AA; 162493 MW; C4FB3001BD09A30D CRC64;

Query Match 30.0%; Score 6; DB 1; Length 1437;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMFL 20
|||
Db 560 LETMFL 565


```

RESULT 23
Q9V9A9
ID Q9V9A9 PRELIMINARY; PRT; 1703 AA.
AC Q9V9A9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE L(2)01289 long form (CG9432-PB).
GN Name=L(2)01289; ORFNames=CG9432;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]_TaxID=7227;
RN SEQUENCE FROM N.A.
RP Manning G., Micklem D.R., Krasnow M.A.;
RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.A., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster K., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heitman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh K.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RN Science 287:2185-2195(2000).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergrén E.J.,
RA Svirkas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RN Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

```

```

RN SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirkas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective."
RN Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
[5]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RN Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
[6]
RN SEQUENCE FROM N.A.
RP FLYBASE;
RN Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[7]
RN SEQUENCE FROM N.A.
RP FLYBASE;
RN Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF216973; AAF34746.1; -.
DR EMBL; AF003790; AAN16126.1; -.
DR HSSP; Q57755; 1F05.
DR IntAct; Q9V9A9; -.
DR FlyBase; FBgn010482; l(2)01289.
DR GO; GO:0005489; P:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006662; Thior.
DR InterPro; IPR006663; Thior.
DR Pfam; PF00085; Thior.
DR Redox-active center.
SQ SEQUENCE 1703 AA; 196579 MW; DAA9C2463426ABDA CRC64;

Query Match 30.0%; Score 6; DB 2; Length 1703;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLETM 18
DB 1348 QMLETM 1353

RESULT 24
Q83XK5 PRELIMINARY; PRT; 23 AA.
ID Q83XK5;
AC Q83XK5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
RN OrderedLocusNames=t2947;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
[1]_TaxID=601;
RN SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18."

```



```
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AB016844; AAO70500.1; -.
KW Hypothetical protein.
SQ SEQUENCE 23 AA; 2797 MW; 9EE0BCB88CB36AA CRC64;

Query Match      25.0%; Score 5; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCOML 15
DB 16 VCOML 20

RESULT 25
Q80KE5 PRELIMINARY; PRT; 28 AA.
AC Q80KE5;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192397; AAO45872.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match      25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
DB 21 PRGAP 25

RESULT 26
Q80KE7 PRELIMINARY; PRT; 28 AA.
AC Q80KE7;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192395; AAO45870.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match      25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 PRGAP 5
DB 21 PRGAP 25

RESULT 27
Q80KE9 PRELIMINARY; PRT; 28 AA.
AC Q80KE9;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192393; AAO45868.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3024 MW; 2E24A7A9AACAC3C6A CRC64;

Query Match      25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
DB 21 PRGAP 25

RESULT 28
Q80KF0 PRELIMINARY; PRT; 28 AA.
AC Q80KF0;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192392; AAO45867.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match      25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
DB 21 PRGAP 25

RESULT 29
Q80KF2 PRELIMINARY; PRT; 28 AA.
AC Q80KF2;
```

DT 01-JUN-2003 (TremBLrel. 24, Created)
 DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
 DE L protein (Fragment).
 OS Rabies virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Lyssavirus.
 OX NCBI_TaxID=11292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22541850; PubMed=12655080;
 RA Paez A., Nunez C., Garcia C., Boshell J.;
 RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
 human and dog rabies associated with bats.";
 RL J. Gen. Virol. 84:795-802(2003).
 DR EMBL; AY192390; AAO45865.1; --
 FT NON_TER 28
 SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 1 PRGAP 5
 Db 21 PRGAP 25

RESULT 30
 Q80KF3
 ID Q80KF3 PRELIMINARY; PRT; 28 AA.
 AC Q80KF3;
 DT 01-JUN-2003 (TremBLrel. 24, Created)
 DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
 DE L protein (Fragment).
 OS Rabies virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Lyssavirus.
 OX NCBI_TaxID=11292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22541850; PubMed=12655080;
 RA Paez A., Nunez C., Garcia C., Boshell J.;
 RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
 human and dog rabies associated with bats.";
 RL J. Gen. Virol. 84:795-802(2003).
 DR EMBL; AY192389; AAO45864.1; --
 FT NON_TER 28
 SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 1 PRGAP 5
 Db 21 PRGAP 25

RESULT 31
 Q80KF4
 ID Q80KF4 PRELIMINARY; PRT; 28 AA.
 AC Q80KF4;
 DT 01-JUN-2003 (TremBLrel. 24, Created)
 DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
 DE L protein (fragment).
 OS Rabies virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Lyssavirus.
 OX NCBI_TaxID=11292;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=22541850; PubMed=12655080;
 RA Paez A., Nunez C., Garcia C., Boshell J.;
 RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
 human and dog rabies associated with bats.";
 RL J. Gen. Virol. 84:795-802(2003).
 DR EMBL; AY192388; AAO45863.1; --
 FT NON_TER 28
 SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 1 PRGAP 5
 Db 21 PRGAP 25

RESULT 32
 Q80KF5
 ID Q80KF5 PRELIMINARY; PRT; 28 AA.
 AC Q80KF5;
 DT 01-JUN-2003 (TremBLrel. 24, Created)
 DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
 DE L protein (Fragment).
 OS Rabies virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Lyssavirus.
 OX NCBI_TaxID=11292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22541850; PubMed=12655080;
 RA Paez A., Nunez C., Garcia C., Boshell J.;
 RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
 human and dog rabies associated with bats.";
 RL J. Gen. Virol. 84:795-802(2003).
 DR EMBL; AY192387; AAO45862.1; --
 FT NON_TER 28
 SQ SEQUENCE 28 AA; 2938 MW; 2B2237A9AB8F3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 1 PRGAP 5
 Db 21 PRGAP 25

RESULT 33
 Q80KF6
 ID Q80KF6 PRELIMINARY; PRT; 28 AA.
 AC Q80KF6;
 DT 01-JUN-2003 (TremBLrel. 24, Created)
 DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
 DE L protein (Fragment).
 OS Rabies virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Lyssavirus.
 OX NCBI_TaxID=11292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22541850; PubMed=12655080;
 RA Paez A., Nunez C., Garcia C., Boshell J.;
 RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
 human and dog rabies associated with bats.";
 RL J. Gen. Virol. 84:795-802(2003).
 DR EMBL; AY192386; AAO45861.1; --
 FT NON_TER 28
 SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 21 PRGAP 25

RESULT 34
Q80KF8 PRELIMINARY; PRT; 28 AA.
AC Q80KF8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192384; AAO45859.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 21 PRGAP 25

RESULT 35
Q80KF9 PRELIMINARY; PRT; 28 AA.
AC Q80KF9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192383; AAO45858.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 21 PRGAP 25

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 21 PRGAP 25

RESULT 36
Q80KG1 PRELIMINARY; PRT; 28 AA.
AC Q80KG1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192381; AAO45856.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 21 PRGAP 25

RESULT 37
Q80KG2 PRELIMINARY; PRT; 28 AA.
AC Q80KG2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192380; AAO45855.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 21 PRGAP 25

RESULT 38
Q80KG3 PRELIMINARY; PRT; 28 AA.
AC Q80KG3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE L protein (Fragment).

```

OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192379; AAO45854.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 39
Q80KG4 PRELIMINARY; PRT; 28 AA.
AC Q80KG4;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192378; AAO45853.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 40
Q80KG7 PRELIMINARY; PRT; 28 AA.
AC Q80KG7;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for

```

```

RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192375; AAO45850.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 2982 MW; 2E2927A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 41
Q80KG8 PRELIMINARY; PRT; 28 AA.
AC Q80KG8;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192374; AAO45849.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 2998 MW; 2E3147A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 42
Q80KG9 PRELIMINARY; PRT; 28 AA.
AC Q80KG9;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192373; AAO45848.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

```



```

RN  SEQUENCE FROM N.A.
RP  MEDLINE=22541850; PubMed=12655080;
RA  Paz A., Nunez C., Garcia C., Boshell J.;
RT  "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT  human and dog rabies associated with bats.";
RL  J. Gen. Virol. 84:795-802(2003).
DR  EMBL; AY192412; AAO31967.1; -.
FT  NON_TER 28
SQ  SEQUENCE 28 AA; 2996 MW; 2F7337A9AACAC3C6A CRC64;

Query Match      25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 PRGAP 5
DB  21 PRGAP 25

RESULT 48
Q80KK4
ID  Q80KK4 PRELIMINARY; PRT; 28 AA.
AC  Q80KK4;
DT  01-JUN-2003 (TReMBLrel. 24, Created)
DT  01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT  01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE  L protein (Fragment).
OS  Rabies virus.
OC  Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC  Rhabdoviridae; Lyssavirus.
OX  NCBI_TaxID=11292;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22541850; PubMed=12655080;
RA  Paz A., Nunez C., Garcia C., Boshell J.;
RT  "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT  human and dog rabies associated with bats.";
RL  J. Gen. Virol. 84:795-802(2003).
DR  EMBL; AY192410; AAO31965.1; -.
FT  NON_TER 28
SQ  SEQUENCE 28 AA; 2996 MW; 2E362C0810CA3C6A CRC64;

Query Match      25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 PRGAP 5
DB  21 PRGAP 25

RESULT 49
Q80KK5
ID  Q80KK5 PRELIMINARY; PRT; 28 AA.
AC  Q80KK5;
DT  01-JUN-2003 (TReMBLrel. 24, Created)
DT  01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT  01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE  L protein (Fragment).
OS  Rabies virus.
OC  Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC  Rhabdoviridae; Lyssavirus.
OX  NCBI_TaxID=11292;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22541850; PubMed=12655080;
RA  Paz A., Nunez C., Garcia C., Boshell J.;
RT  "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT  human and dog rabies associated with bats.";
RL  J. Gen. Virol. 84:795-802(2003).
DR  EMBL; AY192409; AAO31964.1; -.
FT  NON_TER 28
SQ  SEQUENCE 28 AA; 2996 MW; 2F7337A9AACAC3C6A CRC64;

```

```

SQ  SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match      25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 PRGAP 5
DB  21 PRGAP 25

RESULT 50
Q80KK6
ID  Q80KK6 PRELIMINARY; PRT; 28 AA.
AC  Q80KK6;
DT  01-JUN-2003 (TReMBLrel. 24, Created)
DT  01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT  01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE  L protein (Fragment).
OS  Rabies virus.
OC  Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC  Rhabdoviridae; Lyssavirus.
OX  NCBI_TaxID=11292;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22541850; PubMed=12655080;
RA  Paz A., Nunez C., Garcia C., Boshell J.;
RT  "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT  human and dog rabies associated with bats.";
RL  J. Gen. Virol. 84:795-802(2003).
DR  EMBL; AY192408; AAO31963.1; -.
FT  NON_TER 28
SQ  SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match      25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 PRGAP 5
DB  21 PRGAP 25

RESULT 51
Q80KK7
ID  Q80KK7 PRELIMINARY; PRT; 28 AA.
AC  Q80KK7;
DT  01-JUN-2003 (TReMBLrel. 24, Created)
DT  01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT  01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE  L protein (Fragment).
OS  Rabies virus.
OC  Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC  Rhabdoviridae; Lyssavirus.
OX  NCBI_TaxID=11292;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22541850; PubMed=12655080;
RA  Paz A., Nunez C., Garcia C., Boshell J.;
RT  "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT  human and dog rabies associated with bats.";
RL  J. Gen. Virol. 84:795-802(2003).
DR  EMBL; AY192407; AAO31962.1; -.
FT  NON_TER 28
SQ  SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match      25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 PRGAP 5
DB  21 PRGAP 25

```

```
RESULT 52
ID Q80KK8 PRELIMINARY; PRT; 28 AA.
AC Q80KK8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192406; AA031961.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 53
ID Q80KK9 PRELIMINARY; PRT; 28 AA.
AC Q80KK9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192405; AA031960.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 54
ID Q80KLO PRELIMINARY; PRT; 28 AA.
AC Q80KLO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192404; AA031959.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 55
ID Q80KLL PRELIMINARY; PRT; 28 AA.
AC Q80KLL;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192403; AA031958.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 2938 MW; 2F7335C43ACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 56
ID Q80KL2 PRELIMINARY; PRT; 28 AA.
AC Q80KL2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192403; AA031958.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 2938 MW; 2F7335C43ACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25
```

RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
 RL human and dog rabies associated with bats.";
 RL J. Gen. Virol. 84:795-802(2003).
 DR EMBL; AY192402; AAO31957.1; -.
 FT NON_TER 28 28
 SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 Db 21 PRGAP 25

RESULT 57
 Q80KL3 ID Q80KL3 PRELIMINARY; PRT; 28 AA.
 AC Q80KL3;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE L protein (Fragment).
 OS Rabies virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Lyssavirus.
 OX NCBI_TaxID=11292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22541850; PubMed=12655080;
 RA Paez A., Nunez C., Garcia C., Boshell J.;
 RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
 RT human and dog rabies associated with bats.";
 RL J. Gen. Virol. 84:795-802(2003).
 DR EMBL; AY192401; AAO31956.1; -.
 FT NON_TER 28 28
 SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 Db 21 PRGAP 25

RESULT 58
 Q80KL5 ID Q80KL5 PRELIMINARY; PRT; 28 AA.
 AC Q80KL5;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE L protein (Fragment).
 OS Rabies virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Lyssavirus.
 OX NCBI_TaxID=11292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22541850; PubMed=12655080;
 RA Paez A., Nunez C., Garcia C., Boshell J.;
 RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
 RT human and dog rabies associated with bats.";
 RL J. Gen. Virol. 84:795-802(2003).
 DR EMBL; AY192399; AAO31954.1; -.
 FT NON_TER 28 28
 SQ SEQUENCE 28 AA; 2938 MW; 2B2237A9AB8F3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PRGAP 5
 Db 21 PRGAP 25

RESULT 59
 Q9IWC4 ID Q9IWC4 PRELIMINARY; PRT; 28 AA.
 AC Q9IWC4;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE L protein (Fragment).
 OS Rabies virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Lyssavirus.
 OX NCBI_TaxID=11292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22541850; PubMed=12655080;
 RA Paez A., Nunez C., Garcia C., Boshell J.;
 RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
 RT human and dog rabies associated with bats.";
 RL J. Gen. Virol. 84:795-802(2003).
 DR EMBL; AF269292; AAF73515.1; -.
 FT NON_TER 28 28
 SQ SEQUENCE 28 AA; 3024 MW; 2E24A7A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 Db 21 PRGAP 25

RESULT 60
 Q9IWC5 ID Q9IWC5 PRELIMINARY; PRT; 28 AA.
 AC Q9IWC5;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE L protein (Fragment).
 OS Rabies virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Lyssavirus.
 OX NCBI_TaxID=11292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22541850; PubMed=12655080;
 RA Paez A., Nunez C., Garcia C., Boshell J.;
 RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
 RT human and dog rabies associated with bats.";
 RL J. Gen. Virol. 84:795-802(2003).
 DR EMBL; AF269291; AAF73514.1; -.
 DR EIR; PQ0369; PQ0369.
 FT NON_TER 28 28
 SQ SEQUENCE 28 AA; 3024 MW; 2E24A7A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 Db 21 PRGAP 25

RESULT 61


```
Q9QD10
ID Q9QD10 PRELIMINARY; PRT; 28 AA.
AC Q9QD10;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189361; AAF13916.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 62
Q9QD11
ID Q9QD11 PRELIMINARY; PRT; 28 AA.
AC Q9QD11;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189360; AAF13915.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 63
Q9QD12
ID Q9QD12 PRELIMINARY; PRT; 28 AA.
AC Q9QD12;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189359; AAF13914.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 64
Q9QD13
ID Q9QD13 PRELIMINARY; PRT; 28 AA.
AC Q9QD13;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189358; AAF13913.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 65
Q9QD14
ID Q9QD14 PRELIMINARY; PRT; 28 AA.
AC Q9QD14;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189357; AAF13912.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25
```

DR EMBL; AF189357; AAF13912.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRGAP 5
|||||
Db 21 PRGAP 25

RESULT 66
Q9QD15
ID Q9QD15 PRELIMINARY; PRT; 28 AA.
AC Q9QD15;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats."
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189356; AAF13911.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRGAP 5
|||||
Db 21 PRGAP 25

RESULT 67
Q9QD16
ID Q9QD16 PRELIMINARY; PRT; 28 AA.
AC Q9QD16;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats."
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189355; AAF13910.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRGAP 5

Db 21 PRGAP 25
|||||

RESULT 68
Q9QD17
ID Q9QD17 PRELIMINARY; PRT; 28 AA.
AC Q9QD17;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats."
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189354; AAF13909.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3058 MW; 2E2D37A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRGAP 5
|||||
Db 21 PRGAP 25

RESULT 69
Q9QD18
ID Q9QD18 PRELIMINARY; PRT; 28 AA.
AC Q9QD18;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats."
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189353; AAF13908.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3024 MW; 2E24A7A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRGAP 5
|||||
Db 21 PRGAP 25

RESULT 70
Q9QD19
ID Q9QD19 PRELIMINARY; PRT; 28 AA.
AC Q9QD19;
DT 01-MAY-2000 (TREMELrel. 13, Created)

```
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189348; AAF13903.1; -.
FT NON_TER 28 28
SQ SEQUENCE 28 AA; 3024 MW; 2E24A7A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 71
Q9QD22
ID Q9QD22 PRELIMINARY; PRT; 28 AA.
AC Q9QD22;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189349; AAF13904.1; -.
FT NON_TER 28 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 72
Q9QD23
ID Q9QD23 PRELIMINARY; PRT; 28 AA.
AC Q9QD23;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE L protein (fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189348; AAF13903.1; -.
FT NON_TER 28 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 73
Q9QD24
ID Q9QD24 PRELIMINARY; PRT; 28 AA.
AC Q9QD24;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189347; AAF13902.1; -.
FT NON_TER 28 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 74
Q9QD25
ID Q9QD25 PRELIMINARY; PRT; 28 AA.
AC Q9QD25;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189346; AAF13901.1; -.
FT NON_TER 28 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;
```

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred.No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 75
Q9QD26 Q9QD26 PRELIMINARY; PRT; 28 AA.
AC Q9QD26;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats."
PL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AF189345; AAF13900.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACA3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred.No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

Search completed: October 26, 2004, 07:22:24
Job time : 93.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:46:31 ; Search time 85 Seconds
(without alignments)

84.407 Million cell updates/sec

Title: US-10-066-965A-3

Perfect score: 20

Sequence: 1 PRGAPWWRWVQMLETMFL 20

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2002273 seqs, 358729299 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	30.0	34	5	ABP43209 Human ova
2	6	30.0	147	2	AAY76621 Human ova
3	6	30.0	324	2	AAY08348 Murine Oc
4	6	30.0	324	6	ABP97398 Mouse emb
5	6	30.0	358	4	ABG26897 Novel hum
6	6	30.0	501	4	AAG58822 Human GPR
7	6	30.0	731	7	ADJ71063 Human hea
8	6	30.0	1630	4	ABBS59874 Drosophil
9	5	25.0	9	2	AAR37732 Collagen-
10	5	25.0	9	2	AAR93239 Collagen-
11	5	25.0	9	2	AAW57686 Collagen-
12	5	25.0	9	5	AAU71713 Human MHC
13	5	25.0	9	5	AAU71607 Human MHC
14	5	25.0	10	5	AAU71662 Human MHC
15	5	25.0	15	2	AAU93209 New contr
16	5	25.0	20	7	ADC99233 Cancer-re
17	5	25.0	20	8	ADK01531 Hepatitis
18	5	25.0	25	8	ADM32132 EphB4 blo
19	5	25.0	26	2	AAU20230 Human bet
20	5	25.0	28	2	AAU19651 SEQ ID NO
21	5	25.0	33	3	AAG56952 Arabidops
22	5	25.0	42	5	ADF94734 Hepatitis
23	5	25.0	42	5	ADF94732 Hepatitis
24	5	25.0	44	2	AAW57222 Targeting
25	5	25.0	46	3	AAU44951 Xenopus t

26	5	25.0	46	4	AAW21142 Peptide #
27	5	25.0	46	4	AAU15409 Peptide #
28	5	25.0	46	4	ABBA3459 Peptide #
29	5	25.0	46	4	ABBA3414 Peptide #
30	5	25.0	46	4	AAU27897 Peptide #
31	5	25.0	46	4	AAU37347 Peptide #
32	5	25.0	46	4	ABBI19826 Protein #
33	5	25.0	46	4	ABBI19826 Protein #
34	5	25.0	46	4	AAU67599 Human bon
35	5	25.0	46	4	AAU64389 Human bra
36	5	25.0	46	4	ABG58835 Human liv
37	5	25.0	46	4	ABG49246 Human liv
38	5	25.0	46	4	AAU031170 Peptide #
39	5	25.0	46	5	ABG46222 Human pep
40	5	25.0	46	5	ABG37192 Human pep
41	5	25.0	47	4	ABBA38152 Peptide #
42	5	25.0	47	4	AAU31580 Peptide #
43	5	25.0	47	4	ABBA23358 Protein #
44	5	25.0	47	4	AAU71301 Human bon
45	5	25.0	47	4	AAU58788 Human bra
46	5	25.0	47	4	ABG53012 Human liv
47	5	25.0	47	5	ABG41101 Human pep
48	5	25.0	50	4	ABG19401 Novel hum
49	5	25.0	54	3	AAU57910 Arabidops
50	5	25.0	54	3	AAU61712 Arabidops
51	5	25.0	54	6	ABU62915 Mouse ost
52	5	25.0	55	3	ABG60769 Arabidops
53	5	25.0	57	5	ABP64607 Human ORF
54	5	25.0	62	4	AAU46690 Propionib
55	5	25.0	62	6	AAU43209 Propionib
56	5	25.0	63	3	AAU60768 Arabidops
57	5	25.0	63	8	ABO58509 Human gen
58	5	25.0	64	5	ABP00767 Human ORF
59	5	25.0	65	3	AAU61788 Arabidops
60	5	25.0	66	5	ADF94731 Hepatitis
61	5	25.0	67	5	ADF94733 Hepatitis
62	5	25.0	67	4	AAU83752 Human imm
63	5	25.0	67	5	ABP64418 Human ORF
64	5	25.0	68	4	AAU61722 Propionib
65	5	25.0	68	6	ABMS8241 Propionib
66	5	25.0	70	3	AAU57909 Arabidops
67	5	25.0	70	4	AAU10516 Human pol
68	5	25.0	71	4	AAU64459 Propionib
69	5	25.0	71	6	ABM60978 Propionib
70	5	25.0	72	4	AAU14327 Peptide #
71	5	25.0	72	4	ABBA3274 Peptide #
72	5	25.0	72	4	AAU26737 Peptide #
73	5	25.0	72	4	ABBA28100 Human pep
74	5	25.0	72	4	ABBA18736 Protein #
75	5	25.0	72	4	AAU66456 Human bon
76	5	25.0	72	4	AAU54066 Human bra
77	5	25.0	72	4	ABG48123 Human liv
78	5	25.0	72	4	AAU02056 Peptide #
79	5	25.0	72	5	ABU04896 Human pro
80	5	25.0	72	5	ABG36106 Human pep
81	5	25.0	74	4	ABG03203 Novel hum
82	5	25.0	74	4	ABG10885 Novel hum
83	5	25.0	74	5	ABP40006 Staphyloc
84	5	25.0	74	5	ABP38229 Staphyloc
85	5	25.0	76	6	AAU41729 Propionib
86	5	25.0	76	6	ABBA38248 Propionib
87	5	25.0	78	8	AAU09705 Human pol
88	5	25.0	78	8	ADN99367 Novel hum
89	5	25.0	79	4	AAU64243 Propionib
90	5	25.0	79	6	ABM60762 Propionib
91	5	25.0	80	6	ABU11502 Human MDD
92	5	25.0	81	3	AAU60767 Arabidops
93	5	25.0	82	3	AAU27872 Arabidops
94	5	25.0	84	3	AAU27871 Arabidops
95	5	25.0	84	3	AAU27871 Arabidops
96	5	25.0	85	3	AAU28055 Arabidops
97	5	25.0	85	3	AAU28055 Arabidops
98	5	25.0	85	5	ABP64016 Human ORF

99 5 25.0 86 4 AAM18310 Peptide #
100 5 25.0 86 4 AAM70473 Human bon

ALIGNMENTS

RESULT 1
ABP43209
ID ABP43209 standard; protein; 34 AA.
XX AC
XX ABP43209;
DT 22-AUG-2002 (first entry)
XX Human ovarian antigen HVVDQ49, SEQ ID NO:4341.
DE Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive.
XX Homo sapiens.
OS
XX WO200200677-A1.
PN
XX 03-JAN-2002.
PD
XX 07-JUN-2001; 2001WO-US018569.
PF
XX 07-JUN-2000; 2000US-0209467P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Birse CE, Rosen CA;
PI WPI; 2002-147878/19.
XX N-PSDB; ABQ56286.
DR
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT cancer), immune disorders, cardiovascular disorders and neurological
PT diseases.
XX Claim 11; SEQ ID NO 4341; 2922pp; English.
PS
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54311-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may

CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 34 AA;

Query Match 30.0%; Score 6; DB 5; Length 34;
Best Local Similarity 100.0%; Pred.No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMFL 20
Db |||||
2 LETMFL 7

RESULT 2

AY76621
ID AAY76621 standard; protein; 147 AA.

XX AC AAY76621;

DT 10-APR-2000 (first entry)

XX Human ovarian tumor EST fragment encoded protein 117.

XX Expressed sequence tag; EST; human; ovarian tumor; anticancer;
KW gene therapy; treatment.

XX Homo sapiens.

XX DE19817557-A1.

XX 21-OCT-1999.

XX 09-APR-1998; 98DE-01017557.

XX 09-APR-1998; 98DE-01017557.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

XX WPI; 1999-591920/51.

XX N-PSDB; AAZ77497.

XX New nucleic acid sequences expressed in ovarian, and some other, cancer
PT tissues, and derived polypeptides, for treatment of ovarian cancer and
PT identification of therapeutic agents.

XX Claim 25; Page 292; 310pp; German.

XX This invention describes novel nucleic acid (cDNA) sequences (A) which
CC have anticancer activity and are highly expressed in ovarian tumor tissue
CC (and some also in testis and breast cancer tissue). The products of the
CC invention can be used for gene therapy. (A) are used (i) for recombinant
CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)
CC are used (i) to identify agents suitable for treatment of ovarian cancer;
CC (ii) directly for treating this form of cancer (including expression from
CC gene therapy vectors) and (iii) for generation of specific antibodies.
CC (A) are identified by assembling ESTs (expressed sequence tags) from a
CC particular tissue type before comparison of expression patterns. This
CC allows a significantly longer fragment of the gene to be revealed, so
CC should reduce the number of failures associated with the fact that ESTs
CC from different libraries may represent different parts of the same
CC unknown gene, distorting the estimated frequency of occurrence in a
CC particular tissue. AAY76505-Y76638 represent protein fragments encoded by
CC the human ovarian tumor cDNA library derived EST fragments encoded in
CC AAZ77450-Z77572

XX SQ Sequence 147 AA; Query Match 30.0%; Score 6; DB 2; Length 147; Best Local Similarity 100.0%; Pred. No. 50; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAPM 6
|||||

Db 27 PRGAPM 32

RESULT 3
AA08348
ID AAY08348 standard; protein; 324 AA.
XX AC AAY08348;
XX DT 26-JUL-1999 (first entry)
XX DE Murine Oct-4 protein.
XX KW Oct-4; porcine; selectable marker construct; promoter; stem cell;
XX KW differential expression; selective isolation; pluripotent; propagation;
XX KW cell surface membrane protein; rejection; xenotransplantation; murine;
XX KW transgenic pig; organ; transplant; programmed cell death.
XX OS Mus sp.
XX PN WO9919459-A1.
XX PD 22-APR-1999.
XX PF 09-OCT-1998; 98WO-US021289.
XX PR 09-OCT-1997; 97US-00948113.
XX PA (BIOT-) BIOTRANSPLANT INC.
XX PI Baetscher MW, Akiyoshi DE, Kaplan RA;
XX DR WPI; 1999-312473/26.
XX DR N-PSDB; AAX57201.
XX PT Porcine stem cells comprising a marker gene under an Oct-4 promoter.
XX PS Disclosure; Fig 2A-C; 69pp; English.
XX CC This sequence describes novel methods where a genetic selectable marker
CC construct operatively linked to a porcine promoter is introduced into a
CC source of cells that contain porcine stem cells under conditions that
CC allows for differential expression of the marker and hence the selective
CC isolation and/or propagation of desired porcine stem cells. The methods
CC are used to isolate and/or enrich and/or selectively propagate
CC pluripotent porcine cells and to genetically modify porcine cells. The
CC embryonic stem cells could be altered so as to not express a cell surface
CC membrane protein that may cause rejection of porcine cells after
CC xenotransplantation. Transgenic pigs generated through the modified
CC porcine stem cells are useful for providing organs suitable for
CC transplants. The methods overcome the problem of prior art methods
CC associated with the continuing presence of certain differentiated cell
CC types that can cause elimination of stem cells from the culture by
CC inducing their differentiation of programmed cell death. The invention
CC specifically describes the use of the porcine Oct-4 promoter
CC polynucleotide sequence

XX SQ Sequence 324 AA; Query Match 30.0%; Score 6; DB 2; Length 324; Best Local Similarity 100.0%; Pred. No. 96; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMFL 20
|||||

Db 210 LETMFL 215
|||||

RESULT 4
ABP97398
ID ABP97398 standard; protein; 324 AA.
XX AC ABP97398;
XX DT 16-MAY-2003 (first entry)
XX DE Mouse embryonic stem cell Oct-3/4 protein, SEQ ID NO:26.
XX KW Mouse; murine; ES cell associated transcript; ECAT; embryonic stem cell;
XX KW ESC; screening; expression analysis; cell therapy; regeneration therapy;
XX KW tissue regeneration; heart muscle; nerve cell; myocardial infarction;
XX KW neurodegenerative disease; cardiac; neuroprotective; Oct-3/4.
XX OS Mus musculus.
XX PN WO200297090-A1.
XX PD 05-DEC-2002.
XX PF 31-MAY-2002; 2002WO-JP005350.
XX PR 31-MAY-2001; 2001JP-00165927.
XX PA (SUMU) SUMITOMO PHARM CO LTD.
XX PI (YAMA/) YAMANAKA S.
XX PI Yamanaka S, Kaiho H;
XX DR WPI; 2003-140475/13.
XX DR N-PSDB; ABZ75095.
XX PT Screening probes for identifying genes with embryonic stem cell specific
XX PT expression for efficient production of stem cells as regeneration therapy
XX PT treatment of heart and neurological disorders.
XX PS Claim 29; Page 78-79; 125pp; Japanese.
XX CC The invention relates to probes for screening embryonic stem (ES) cells.
XX CC The probes contain polynucleotide sequences from genes designated ES cell
XX CC associated transcript (ECAT) genes, which are specifically expressed in
XX CC ES cells. The polynucleotide sequences of 16 murine and human ECAT genes
XX CC (ABZ75087-ABZ75094, ABZ75096-ABZ75101, ABZ75103-ABZ75104) are fully
XX CC defined in the specification, as are their encoded proteins (ABP97390-
XX CC ABP97397, ABP97399-ABP97404, ABP97406-ABP97407). The probes of the
XX CC invention are used to analyse the expression of ECAT genes in ES cells,
XX CC providing an efficient means of screening ES cells in order to obtain
XX CC large amounts of ES cells. The ES cells thus obtained can be used in
XX CC regeneration therapy of heart muscle and nerve cells for treatment of
XX CC myocardial infarction and degenerative diseases of the nervous system.
XX CC The present sequence represents murine Oct-3/4 protein which is
XX CC specifically claimed for use in the invention

XX SQ Sequence 324 AA; Query Match 30.0%; Score 6; DB 6; Length 324; Best Local Similarity 100.0%; Pred. No. 96; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMFL 20
|||||

Db 210 LETMFL 215
|||||

RESULT 5
ABG26897
ID ABG26897 standard; protein; 358 AA.
XX

AC ABG26897;
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #26888.
 XX
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 KW
 XX
 OS Homo sapiens.
 XX
 XX WO200175067-A2.
 PN
 XX
 PD 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US008631.
 PF
 XX 31-MAR-2000; 2000US-00540217.
 XX PR
 XX 23-AUG-2000; 2000US-00649167.
 XX PR
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 PI
 XX
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS91084.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 PT
 XX
 XX Claim 20; SEQ ID NO 57256; 103pp; English.
 PS
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 358 AA;
 SQ
 Query Match 30.0%; Score 6; DB 4; Length 358;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 15 LETMFL 20
 DB 39 LETMFL 44
 |||||
 RESULT 6
 AAG65822
 ID AAG65822 standard; protein; 501 AA.
 XX
 AC AAG65822;
 XX

DT 30-JAN-2002 (first entry)
 XX
 DE Human GPR38 variant GPR38V polypeptide.
 XX
 KW GPR38V; variant; antibacterial; cytostatic; analgesic; antiasthmatic;
 KW anti-Parkinsonian; hypertensive; hypotensive; antidiabetic; osteopathic;
 KW anti-allergic; antimigraine; neuroleptic; nootropic; anticonvulsant;
 KW antiulcer; antiemetic; cardiant; vaccine; human.
 XX
 OS Homo sapiens.
 XX
 XX WO200164836-A2.
 PN
 XX 07-SEP-2001.
 PD
 XX 28-FEB-2001; 2001WO-US006277.
 PF
 XX 01-MAR-2000; 2000US-00516315.
 XX PR
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA
 XX Elshourbagy N, Shabon U;
 PI
 XX WPI; 2001-639956/73.
 DR N-PSDB; AAT66989.
 XX
 XX New human GPR38V polypeptide and polynucleotide, useful for treating e.g.
 PT bacterial, fungal, protozoal and viral infections, cancers or allergies,
 PT as vaccines, and for identifying agonists and antagonists potentially
 PT useful in therapy.
 PT
 XX Claim 1; Page 26; 32pp; English.
 PS
 XX This represents a human GPR38 variant (GPR38V) polypeptide. GPR38V can be
 CC expressed by standard recombinant methodology. The polynucleotides and
 CC polypeptides are used in the treatment of bacterial, fungal, protozoal
 CC and viral infections, infections caused by HIV-1 or HIV-2, pain, cancers,
 CC diabetes, obesity, anorexia, asthma, Parkinson's disease, acute heart
 CC failure, hypertension, urinary retentions, osteoporosis, allergies,
 CC ulcers, migraine, psychotic and neurological disorders, or dyskinesias.
 CC They are also useful for identifying agonists and antagonists that are
 CC potentially useful in therapy, as vaccines to induce immunological
 CC response in a mammal. The polypeptides may also be used as immunogens to
 CC produce antibodies immunospecific for the polypeptides, and to identify
 CC membrane bound or soluble receptors
 XX
 XX Sequence 501 AA;
 SQ
 Query Match 30.0%; Score 6; DB 4; Length 501;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PRGAPM 6
 DB 85 PRGAPM 90
 |||||
 RESULT 7
 ADJ71063
 ID ADJ71063 standard; protein; 731 AA.
 XX
 AC ADJ71063;
 XX
 XX 06-MAY-2004 (first entry)
 DT
 XX Human heat mitochondrial protein as a therapeutic target SeqID2869.
 DE
 XX mitochondrial; human; screening assay; diabetes mellitus;
 KW Huntington's disease; osteoarthritis;
 KW Leber's hereditary optic neuropathy; LHON;
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
 XX

KW osteopathic; ophthalmological; cytostatic.
 XX Homo sapiens.
 OS
 PN WO2003087768-A2.
 XX
 XX 23-OCT-2003.
 PD
 PF 04-APR-2003; 2003WO-US010870.
 XX
 XX 12-APR-2002; 2002US-0372843P.
 PR 17-JUN-2002; 2002US-0389987P.
 PR 20-SEP-2002; 2002US-0412418P.
 XX
 XX (MITO-) MITOKOR.
 PA (BUCK-) BUCK INST AGE RES.
 PA
 PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
 PI Warnock DE;
 XX
 XX WPI; 2003-845369/78.
 DR
 XX
 XX Identifying a mitochondrial target for drug screening assays and for
 PT treating diseases associated with altered mitochondrial function,
 PT comprises detecting a modified polypeptide in a sample and correlating
 PT with the disease.
 XX
 XX Claim 1; SEQ ID NO 2859; 180pp; English.
 PS
 XX This invention relates to novel mitochondrial targets that can be used
 CC for therapeutic intervention in treating a disease associated with
 CC altered mitochondrial function. Specifically, it refers to a method for
 CC identifying proteins of the human heart mitochondrial proteome that are
 CC useful for drug screening assays, as well as therapeutic targets. The
 CC present invention describes a method for identifying such proteins that
 CC can be used in the treatment of various diseases associated with altered
 CC mitochondrial function including diabetes mellitus, Huntington's disease,
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, nootropic, antidiabetic,
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 CC cytotstatic activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.
 XX
 XX Sequence 731 AA;
 SQ
 Query Match 30.0%; Score 6; DB 7; Length 731;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PRGAPM 6
 DB 542 PRGAPM 547
 |||||
 RESULT 8
 ABB59874
 ID ABB59874 standard; protein; 1630 AA.
 XX
 XX ABB59874;
 AC
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polypeptide SEQ ID NO 6414.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX

PD 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US009231.
 XX
 XX 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR N-PSDB; ABL03977.
 DR
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 PT
 XX Disclosure; SEQ ID NO 6414; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 1630 AA;
 SQ
 Query Match 30.0%; Score 6; DB 4; Length 1630;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 13 QMLETM 18
 DB 1254 QMLETM 1259
 |||||
 RESULT 9
 AAR37732
 ID AAR37732 standard; peptide; 9 AA.
 XX
 XX AAR37732;
 AC
 XX 25-MAR-2003 (revised)
 DT 07-SEP-1993 (first entry)
 DT
 XX Collagen-like polymer #15.
 DE
 XX Recombinant; collagen-like polymer; CLP; tripeptide; helix; membrane;
 KW fibre; film; coating; triad sequence; collagen; mammalian; moulding;
 KW hydrogel; interchain linkage; colloid suspension.
 XX
 XX Synthetic.
 OS
 XX WO9310154-A1.
 PN
 XX 27-MAY-1993.
 PD
 XX 04-NOV-1992; 92WO-US009485.
 PF
 XX 12-NOV-1991; 91US-00791960.
 PR
 XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
 PA
 XX Cappello J, Ferrari FA;
 PI WPI; 1993-182496/22.
 XX
 XX

PT High mol. wt. collagen-like protein polymers - capable of being produced
 PT in unicellular microorganisms.

XX Disclosure; Page 12; 82pp; English.

XX The sequences given in AAR37718-32 are examples of recombinantly produced
 CC collagen-like polymers (CLPs) which consist of repeated tripeptide
 CC sequences selected from a wide range of GXY sequences, where X and Y can
 CC be any amino acid. These polymers have molecular weights of >30 kD and
 CC are able to form helices due to interchain linkages. These polymers pref.
 CC contain a proportion of tripeptide triad sequences found in natural
 CC collagens, pref. mammalian collagens. The CLPs impart unique
 CC characteristics to materials such as fibres, membranes, films, coatings,
 CC hydrogels, colloid suspensions and moulded articles. (Updated on 25-MAR-
 CC 2003 to correct PN field.)

XX SQ Sequence 9 AA;

Query Match 25.0%; Score 5; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 |||||
 DB 5 PRGAP 9

RESULT 10

AAR93239
 ID AAR93239 standard; peptide; 9 AA.

XX AC AAR93239;

XX 25-MAR-2003 (revised)

DT 24-FEB-1996 (first entry)

XX Collagen-like GUB sequence motif 5.

XX collagen; repetitive triad motif; recombinant production; photographic;
 KW medical; structural; fibre.

XX Synthetic.

XX US5496712-A.

XX 05-MAR-1996.

XX 05-NOV-1992; 92US-00972032.

XX 06-NOV-1990; 90US-00609716.

PR 12-NOV-1991; 91US-00791960.

XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.

XX Cappelletto J, Ferrari FA;

XX WPI; 1996-150728/15.

XX Collagen-like polymers comprising repetitive triads - produced in
 PT unicellular organisms with improved characteristics, useful in, e.g.
 PT photographic and medical fibres.

XX Disclosure; Col 6; 43pp; English.

XX The invention concerns collagen-like polymers having repetitive triads
 CC with reduced proline content, and where glycine is the initial amino acid
 CC and the subsequent amino acids are varied. The choice of triads utilised
 CC in a recombinant collagen-like polymer are chosen in order to affect
 CC properties such as helix stability, hydration, solubility, gel point,
 CC biodegradation and immunogenicity. Also considered is the level of
 CC guanidine and cytosine nucleotides (due to levels of glycine and proline)
 CC present in the genes encoding the polymers. As the gene is synthesised
 CC there is opportunity for strands to loop out, single- stranded DNA to be

CC excised, recombination events to occur which can result in loss of
 CC segments of the gene, and inefficient transcription and/or translation
 CC (due to the presence of self-complementary sequences), hence genes of the
 CC invention are designed to provide the advantageous properties of
 CC collagen, while at the same time allowing for stable expression of high
 CC mol. wt. collagen-like proteins. Triads of particular interest include
 CC GDR, GSP, GDA, GAH and GEA. The collagen-like polymers may impart new
 CC characteristics, finding wide used in photographic, medical, structural
 CC and fibre applications, and are capable of being produced in unicellular
 CC microorganisms at high mol. wts. and in high efficiency. AAR93235-39 are
 CC peptides defining GUB (sic) sequences. The peptides are useful as
 CC happens, to produce antisera or monoclonal antibodies specific to the
 CC sequences which are then used for affinity purification, identification
 CC of the polymers, etc. . (Updated on 25-MAR-2003 to correct PF field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 9 AA;

Query Match 25.0%; Score 5; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 |||||
 DB 5 PRGAP 9

RESULT 11

AAW57686

ID AAW57686 standard; peptide; 9 AA.

XX AC AAW57686;

XX 27-AUG-1998 (first entry)

XX Collagen-like polymer fragment.

XX Collagen-like polymer; synthetic polymer; fibre coating;
 KW prosthetic device; catalytic substance.

XX Synthetic.

XX US5773249-A.

XX 30-JUN-1998.

XX 02-MAY-1996; 96US-00642255.

XX 04-NOV-1986; 86US-00927258.

PR 29-OCT-1987; 87US-00114618.

PR 09-NOV-1988; 88US-00269429.

PR 06-NOV-1990; 90US-00609716.

PR 12-NOV-1991; 91US-00791960.

PR 05-NOV-1992; 92US-00972032.

PR 22-DEC-1995; 95US-00577046.

XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.

XX Ferrari FA, Cappelletto J;

XX WPI; 1998-387004/33.

XX Recombinant collagen-like polymers - useful for making gels, films,
 PT fibres, etc.

XX Disclosure; Col 6; 93pp; English.

XX This sequence represents a fragment of a unnatural collagen-like polymer
 CC of the invention. The products may be used as films, fibres, moulded
 CC objects and admixed with other natural or synthetic polymers or coatings
 CC on fibres, films, labware or other surfaces, e.g. prosthetic devices. The
 CC polymers may be used for binding a wide variety of specific binding

CC materials, as catalytic substances (where the amino acid sequence may
CC specifically chelate a wide variety of elements), as purification media,
CC composites, laminates or adhesives. They may also be combined with
CC inorganic or organic materials such as carbon fibres, nylon fibres,
CC nitrocellulose, etc., as flask coatings or in synthetic matrices for the
CC growth and study of cells, as affinity columns or as supports for
CC biological materials. The polymers have collagen-like properties, but may
CC be easily expressed in micro-organisms in high efficiency. The new
CC sequences can be tailored to give the desired properties

XX SQ Sequence 9 AA;

Query Match 25.0%; Score 5; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
DB 5 PRGAP 9
|||||

RESULT 12
AAU71713
ID AAU71713 standard; peptide; 9 AA.

XX AAU71713;

AC 26-FEB-2002 (first entry)

DT Human MHC molecule HLA-B35 binding 103P3E8 peptide #20.

DE 103P3E8; prostate; bladder; kidney; colon; lung; breast; rectum; stomach;
KW tumour; cancer; cytostatic; gene therapy; antibody therapy; ribozyme;
KW single chain monoclonal antibody; serum; blood; urine; tissue; human;
KW chromosome 9q13-q21.

XX Homo sapiens.

OS WO200179557-A2.

PN 25-OCT-2001.

PD 12-APR-2001; 2001WO-US012181.

PF 12-APR-2000; 2000US-0196647P.

PR (UROG-) UROGENESYS INC.

PA Faris M, Challita-Eid PM, Raitano AB, Mitchell SC, Afar DEH;
PI Jakobovits A;

XX WPI; 2002-061976/08.

DR Monitoring 103P3E8 gene products in sample from patient (suspected of)
PT having cancer, useful for diagnosing, managing or treating cancers, e.g.
PT prostate cancer, comprises determining presence of aberrant 103P3E8 gene
PT products.

XX Disclosure; Page 98; 128pp; English.

CC Sequences AAU71093-AAU71796 represent the 103P3E8-related protein and
CC peptide fragments of the protein. 103P3E8 exhibits tissue specific
CC expression in normal adult tissue, but it is also aberrantly expressed in
CC many cancers including tumours of the prostate, bladder, kidney, colon,
CC lung, breast, rectum and stomach. The 103P3E8 polynucleotide, its related
CC protein and peptide fragments and specific PCR primers are therefore
CC useful for diagnosing and treating cancer. A vector comprising a
CC polynucleotide which encodes a single chain monoclonal antibody, that
CC immunospecifically binds to an 103P3E8-related protein, and a ribozyme
CC capable of cleaving a polynucleotide having the 103P3E8 coding sequence,
CC are both useful in the preparation of a composition for treating a
CC patient with a cancer that expresses 103P3E8. The sequences can be used
CC in diagnostic methods to monitor the level of 103P3E8 gene products in

CC serum, blood, urine and tissue and to thereby detect the presence of
CC cancerous cells
XX Sequence 9 AA;

Query Match 25.0%; Score 5; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
DB 2 PRGAP 6
|||||

RESULT 13
AAU71607
ID AAU71607 standard; peptide; 9 AA.

XX AAU71607;

AC 26-FEB-2002 (first entry)

DT Human MHC class I molecule HLA-B7 binding 103P3E8 peptide #14.

DE 103P3E8; prostate; bladder; kidney; colon; lung; breast; rectum; stomach;
KW tumour; cancer; cytostatic; gene therapy; antibody therapy; ribozyme;
KW single chain monoclonal antibody; serum; blood; urine; tissue; human;
KW chromosome 9q13-q21.

XX Homo sapiens.

OS WO200179557-A2.

PN 25-OCT-2001.

PD 12-APR-2001; 2001WO-US012181.

PF 12-APR-2000; 2000US-0196647P.

PR (UROG-) UROGENESYS INC.

PA Faris M, Challita-Eid PM, Raitano AB, Mitchell SC, Afar DEH;
PI Jakobovits A;

XX WPI; 2002-061976/08.

DR Monitoring 103P3E8 gene products in sample from patient (suspected of)
PT having cancer, useful for diagnosing, managing or treating cancers, e.g.
PT prostate cancer, comprises determining presence of aberrant 103P3E8 gene
PT products.

XX Disclosure; Page 95; 128pp; English.

CC Sequences AAU71093-AAU71796 represent the 103P3E8-related protein and
CC peptide fragments of the protein. 103P3E8 exhibits tissue specific
CC expression in normal adult tissue, but it is also aberrantly expressed in
CC many cancers including tumours of the prostate, bladder, kidney, colon,
CC lung, breast, rectum and stomach. The 103P3E8 polynucleotide, its related
CC protein and peptide fragments and specific PCR primers are therefore
CC useful for diagnosing and treating cancer. A vector comprising a
CC polynucleotide which encodes a single chain monoclonal antibody, that
CC immunospecifically binds to an 103P3E8-related protein, and a ribozyme
CC capable of cleaving a polynucleotide having the 103P3E8 coding sequence,
CC are both useful in the preparation of a composition for treating a
CC patient with a cancer that expresses 103P3E8. The sequences can be used
CC in diagnostic methods to monitor the level of 103P3E8 gene products in
CC serum, blood, urine and tissue and to thereby detect the presence of
CC cancerous cells

XX SQ Sequence 9 AA;

Query Match 25.0%; Score 5; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;

```
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PRGAP 5
   |||||
Db 2 PRGAP 6

RESULT 14
AAU71662
ID AAU71662 standard; peptide; 10 AA.
XX
AC AAU71662;
XX
DT 26-FEB-2002 (first entry)
XX
DE Human MHC class I molecule HLA-B7 binding 103P3E8 peptide #69.
XX
KW 103P3E8; prostate; bladder; kidney; colon; lung; breast; rectum; stomach;
KW tumour; cancer; cytostatic; gene therapy; antibody therapy; ribozyme;
KW single chain monoclonal antibody; serum; blood; urine; tissue; human;
KW chromosome 9q13-q21.
XX
OS Homo sapiens.
XX
PN WO200179557-A2.
XX
PD 25-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US012181.
XX
PR 12-APR-2000; 2000US-0196647P.
XX
PA (UROG-) UROGENESYS INC.
XX
PI Faris M, Challita-Eid PM, Raitano AB, Mitchell SC, Afar DEH;
PI Jakobovits A;
XX
WPI; 2002-061976/98.
XX
PT Monitoring 103P3E8 gene products in sample from patient (suspected of)
PT having cancer, useful for diagnosing, managing or treating cancers, e.g.
PT prostate cancer, comprises determining presence of aberrant 103P3E8 gene
PT products.
XX
PS Disclosure; Page 96; 128pp; English.
XX
CC Sequences AAU71093-AAU71796 represent the 103P3E8-related protein and
CC peptide fragments of the protein. 103P3E8 exhibits tissue specific
CC expression in normal adult tissue, but it is also aberrantly expressed in
CC many cancers including tumours of the prostate, bladder, kidney, colon,
CC lung, breast, rectum and stomach. The 103P3E8 polynucleotide, its related
CC protein and peptide fragments and specific PCR primers are therefore
CC useful for diagnosing and treating cancer. A vector comprising a
CC polynucleotide which encodes a single chain monoclonal antibody, that
CC immunospecifically binds to an 103P3E8-related protein, and a ribozyme
CC capable of cleaving a polynucleotide having the 103P3E8 coding sequence,
CC are both useful in the preparation of a composition for treating a
CC patient with a cancer that expresses 103P3E8. The sequences can be used
CC in diagnostic methods to monitor the level of 103P3E8 gene products in
CC serum, blood, urine and tissue and to thereby detect the presence of
CC cancerous cells
XX
SQ Sequence 10 AA;
Query Match 25.0%; Score 5; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PRGAP 5
   |||||
Db 2 PRGAP 6
```

```
RESULT 15
AAR93209
ID AAR93209 standard; peptide; 15 AA.
XX
AC AAR93209;
XX
DT 04-OCT-1996 (first entry)
XX
DE New contraceptive peptide #5 derived from Zona Pellucida ZP3 protein.
XX
KW Zona pellucida; ZP3; vaccine; humoral response; contraception; epitope;
KW pathogenic T cell response; ovary; assay; autoimmune; antibody;
KW passive immunisation.
XX
OS Synthetic.
XX
PN WO9606113-A1.
XX
PD 29-FEB-1996.
XX
PF 18-AUG-1995; 95WO-EP003311.
XX
PR 22-AUG-1994; 94EP-00202392.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Van Duin M, Grootenhuis AJ, Bunschoten EJ;
XX
WPI; 1996-151331/15.
XX
PT Immuno:contraceptive peptide(s) derived from Zona Pellucida protein ZP3 -
PT used to prepare contraceptive vaccine and in assays to measure autoimmune
PT antibodies.
XX
PS Claim 5; Page 31; 43pp; English.
XX
CC Peptides AAR93205-9 are examples of peptides derived from the sequence of
CC the Zona Pellucida protein ZP3 which contain the amino acid sequences
CC AAR93210 or AAR93214-5. The peptides are esp. based on amino acids 23-30
CC of the ZP3 protein. The novel peptides can be used in vaccines to induce
CC a humoral response against the ZP3 protein e.g. for contraception, esp.
CC as they do not raise a pathogenic T cell response since they do not
CC contain T cell epitopes. The novel peptides thus avoid potential ovarian
CC damage caused by some peptides used as vaccines. The peptides are also
CC useful in assays for detecting autoimmune antibodies or for generating
CC antibodies for passive immunisation
XX
SQ Sequence 15 AA;
Query Match 25.0%; Score 5; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GAPMW 7
   |||||
Db 3 GAPMW 7

RESULT 16
ADC99233
ID ADC99233 standard; peptide; 20 AA.
XX
AC ADC99233;
XX
DT 01-JAN-2004 (first entry)
XX
DE Cancer-related DGI-2-binder peptide - SEQ ID 66.
XX
KW cytostatic; cancer; gene therapy; DGI-2; DGI-5; DGI-7; DGI-9; Hras;
KW leptin; VEGF; vascular endothelial growth factor receptor; VEGF-R1;
KW VEGF-R2; VEGF-R3; FLT1; FMS-related tyrosine kinase 1; FLK1; KDR;
KW kinase insert domain protein receptor; EGFR; epidermal growth factor;
KW FGFR1; fibroblast growth factor; Tie-1.
XX
```

```

XX OS Unidentified.
XX FN WO2003035839-A2.
XX PD 01-MAY-2003.
XX PF 24-OCT-2002; 2002WO-US034021.
XX PR 24-OCT-2001; 2001US-0345471P.
XX PA (DGIB-) DGI BIOTECHNOLOGIES INC.
XX PI Pillutla RC, Brissette R, Spruyt M, Dedova O, Blume A;
XX PI Prendergast J, Goldstein N;
XX WPI; 2003-457332/43.
XX
XX Selecting target and target binder pairs for preparing a composition for
XX treating cancer by mixing in a reaction vessel phage expressing
XX biological targets and phage expressing target binders.
XX
XX Claim 26; SEQ ID NO 66; 172pp; English.
XX
XX The invention relates to a novel method of selecting target and target
XX binder pairs comprising mixing in a reaction vessel phage expressing
XX biological targets and phage expressing target binders, each having
XX distinguishable selection markers and selecting target and target binder
XX pairs based on the selection markers. The molecules of the invention
XX demonstrate cytostatic activity whilst the method may be useful for
XX selecting target and target binder pairs for preparing a composition for
XX treating cancer. Furthermore, the method may be utilised during gene
XX therapy procedures. The current sequence is that of the cancer-related
XX DGI-2-binder peptide of the invention.
XX
XX Sequence 20 AA;
XX
Query Match 25.0%; Score 5; DB 7; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 6 MMRW 10
Db |||||
8 MMRW 12

RESULT 17
ADK01531
ID ADK01531 standard; peptide; 20 AA.
XX
XX AC ADK01531;
XX
XX 06-MAY-2004 (first entry)
XX
XX Hepatitis C virus 1b peptide SeqID183.
XX
XX pathogenic virus; alternative reading frame; antigenic determinant;
XX virucide; vaccine; therapeutic agent; infection.
XX
XX Hepatitis C virus.
XX
XX WO2004011650-A2.
XX
XX 05-FEB-2004.
XX
XX 24-JUL-2003; 2003WO-EP008112.
XX
XX 24-JUL-2002; 2002AT-00001124.
XX
XX 11-JUL-2003; 2003EP-00450171.
XX
XX (INTR-) INTERCELL AG.
XX
XX Mattner F, Schmidt W, Habel A;

```

```

XX WPI; 2004-159243/16.
XX
XX New polypeptide encoded by an alternative reading frame of a pathogenic
XX virus comprising an antigenic determinant, useful for treating or
XX preventing an infection with the pathogenic virus.
XX
XX Claim 7; SEQ ID NO 183; 220pp; English.
XX
XX This invention relates to a novel polypeptide encoded by an alternative
XX reading frame of a pathogenic virus, where the polypeptide starts with a
XX methionine amino acid residue, which comprises an antigenic determinant
XX and more than 7 amino acid residues. The invention may be useful for the
XX production of compounds with a virucide activity or the development of a
XX vaccine. The polypeptide or its fragments may be useful as a therapeutic
XX agent. It is also useful for the manufacture of a medicament for treating
XX or preventing an infection with the pathogenic virus. The present
XX sequence is that of a hepatitis C virus peptide of the invention.
XX
XX Sequence 20 AA;
XX
Query Match 25.0%; Score 5; DB 8; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 PRGAP 5
Db |||||
7 PRGAP 11

RESULT 18
ADM32132
ID ADM32132 standard; protein; 25 AA.
XX
XX AC ADM32132;
XX
XX 17-JUN-2004 (first entry)
XX
XX EphB4 blocking peptide 6, seq id 7.
XX
XX Cytostatic; gene therapy; Inhibitor; cancerous growth; cancer; breast;
XX prostate; ephrin type-B receptor 4 precursor; EphB4; human.
XX
XX Homo sapiens.
XX
XX WO2004024773-A1.
XX
XX 25-MAR-2004.
XX
XX 16-SEP-2003; 2003WO-AU001209.
XX
XX 16-SEP-2002; 2002AU-00951409.
XX
XX (QUEE-) QUEEN ELIZABETH HOSPITAL.
XX
XX Stephenson S;
XX
XX WPI; 2004-270012/25.
XX
XX Inhibiting cancerous growth of a cell by contacting the cell with an
XX antibody, or its antigen-binding portion, that binds to an epitope of
XX EphB4 polypeptide.
XX
XX Example 6; SEQ ID NO 7; 73pp; English.
XX
XX The invention relates to a method of inhibiting cancerous growth of a
XX cell, comprising contacting the cell with at least one antibody, or its
XX antigen-binding portion, where the antibody or antigen-binding portion
XX binds to an epitope located within residues 200-400 of EphB4, which
XX consists of the sequence of 984 amino acids (S1) fully defined in the
XX specification, or to an epitope located in a sequence at least 85%
XX identical to residues 200-400, 201-245, 220-244 or 220-230 of EphB4. The
XX composition and methods are useful for treating or preventing cancer

```

CC (e.g. breast or prostate cancer) or for identifying agents that inhibit
 CC cancerous growth of a cell. The current sequence represents an EphB4
 CC blocking peptide used in an example from the invention in EphB4 epitope
 CC mapping.

SQ Sequence 25 AA;

Query Match 25.0%; Score 5; DB 8; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 |||||
 Db 15 PRGAP 19

RESULT 19
 AAY20230
 ID AAY20230 standard; protein; 26 AA.

XX AAY20230;

XX 22-JUL-1999 (first entry)

XX Human beta-amyloid precursor protein mutant fragment 55.

XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW Scl-2; B-cell leukemia/lymphoma 2 proto-oncogene; BMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.

OS Homo sapiens.

XX WO9845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-IB000705.

XX 10-APR-1997; 97US-0043163P.

XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

PA (UYRO-) UNIV ROTTERDAM ERASMUS.

PA (UYUT-) RIJKSUNIV UTRECHT.

XX Van Leeuwen FW, Grosveld FG, Burbach JPH;

XX WPI; 1998-609901/51.

DR N-PSDB; AAX75753.

XX Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also for
 PT treatment and prevention with specific ribozymes or wild-type RNA.

XX Disclosure; Fig 2; 25pp; English.

XX This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be

CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (BMGP-C) and neuroendocrine specific protein A

XX Sequence 26 AA;

Query Match 25.0%; Score 5; DB 2; Length 26;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 |||||
 Db 5 PRGAP 9

RESULT 20

AAY19651

ID AAY19651 standard; protein; 28 AA.

XX AAY19651;

XX 14-JUL-1999 (first entry)

DE SEQ ID NO 369 from WO9922243.

XX Human secreted protein; cancer; tumour; neurodegenerative disorder;
 KW developmental abnormality; fetal deficiency; blood disorder; leukemia;
 KW immune system disease; autoimmune disease; hepatic disease;
 KW renal disease; lymphoma; inflammation; allergy; ischemic shock;
 KW Alzheimer's; cognitive disorder; schizophrenia; prostate disease;
 KW obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease;
 KW lung disease; thymus disease; digestive disorder; endocrine disorder;
 KW infection; AIDS.

XX Homo sapiens.

XX WO9922243-A1.

XX 06-MAY-1999.

XX 23-OCT-1998; 98WO-US022376.

XX 24-OCT-1997; 97US-0062784P.

PR 24-OCT-1997; 97US-0063088P.

PR 24-OCT-1997; 97US-0063089P.

PR 24-OCT-1997; 97US-0063090P.

PR 24-OCT-1997; 97US-0063091P.

PR 24-OCT-1997; 97US-0063092P.

PR 24-OCT-1997; 97US-0063097P.

PR 24-OCT-1997; 97US-0063098P.

PR 24-OCT-1997; 97US-0063099P.

PR 24-OCT-1997; 97US-0063100P.

PR 24-OCT-1997; 97US-0063101P.

PR 24-OCT-1997; 97US-0063109P.

PR 24-OCT-1997; 97US-0063110P.

PR 24-OCT-1997; 97US-0063111P.

PR 24-OCT-1997; 97US-0063148P.

PR 24-OCT-1997; 97US-0063386P.

PR 24-OCT-1997; 97US-0063387P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Feng P, Rosen CA, Ruben SM, Ni J, Wei Y, Soppet DR, Moore PA;

PI Kayw H, Lafleur DW, Olsen HS, Brewer LA, Shi Y, Ebner R, Young P;

PI Greene JM, Florence KA, Florence C, Duan DR, Janat F, Endress GA;

PI Carter KC;

XX

DR WPI; 1999-303069/25.
XX PT New isolated human genes and the secreted polypeptides they encode.
XX PS Disclosure; Page 485; 546pp; English.
XX
CC The specification describes human secreted proteins. The polynucleotides
CC and their corresponding secreted polypeptides are useful for preventing,
CC treating or ameliorating medical conditions, e.g. by protein or gene
CC therapy. Pathological conditions can also be diagnosed by determining the
CC amount of the polypeptides in a sample or by determining the presence of
CC mutations in the polynucleotides. Specific uses are described for each of
CC the polynucleotides, based on which tissues they are most highly
CC expressed in, and include developing products for the diagnosis or
CC treatment of cancer, tumours, neurodegenerative disorders, developmental
CC abnormalities and fetal deficiencies, blood disorders, leukemias,
CC diseases of the immune system, autoimmune diseases, hepatic and renal
CC disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's
CC and cognitive disorders, schizophrenia, prostate diseases, obesity,
CC disorders involving osteoclasts such as osteoporosis, arthritis or
CC malignancies, diseases of testes, lung or thymus, digestive/endocrine
CC disorders, infections and AIDS. The polypeptides are also useful for
CC identifying their binding partners
XX
SQ Sequence 28 AA;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CQMLE 16
|||||
Db 3 CQMLE 7

RESULT 21
AAG56952
ID AAG56952 standard; protein; 33 AA.
XX
AC AAG56952;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 73318.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.

30-APR-1999; 99US-0132407P.
04-MAY-1999; 99US-0132484P.
05-MAY-1999; 99US-0132485P.
06-MAY-1999; 99US-0132486P.
06-MAY-1999; 99US-0132487P.
07-MAY-1999; 99US-0132863P.
11-MAY-1999; 99US-0134256P.
14-MAY-1999; 99US-0134218P.
14-MAY-1999; 99US-0134219P.
14-MAY-1999; 99US-0134221P.
14-MAY-1999; 99US-0134370P.
18-MAY-1999; 99US-0134768P.
19-MAY-1999; 99US-0134941P.
20-MAY-1999; 99US-0135124P.
21-MAY-1999; 99US-0135353P.
24-MAY-1999; 99US-0135629P.
25-MAY-1999; 99US-0136021P.
27-MAY-1999; 99US-0136392P.
28-MAY-1999; 99US-0136782P.
01-JUN-1999; 99US-0137222P.
03-JUN-1999; 99US-0137528P.
04-JUN-1999; 99US-0137502P.
07-JUN-1999; 99US-0137724P.
08-JUN-1999; 99US-0138094P.
10-JUN-1999; 99US-0138540P.
10-JUN-1999; 99US-0138847P.
14-JUN-1999; 99US-0139119P.
16-JUN-1999; 99US-0139452P.
16-JUN-1999; 99US-0139453P.
17-JUN-1999; 99US-0139492P.
18-JUN-1999; 99US-0139454P.
18-JUN-1999; 99US-0139455P.
18-JUN-1999; 99US-0139456P.
18-JUN-1999; 99US-0139457P.
18-JUN-1999; 99US-0139458P.
18-JUN-1999; 99US-0139459P.
18-JUN-1999; 99US-0139460P.
18-JUN-1999; 99US-0139461P.
18-JUN-1999; 99US-0139462P.
18-JUN-1999; 99US-0139463P.
18-JUN-1999; 99US-0139750P.
18-JUN-1999; 99US-0139763P.
21-JUN-1999; 99US-0139817P.
22-JUN-1999; 99US-0139899P.
23-JUN-1999; 99US-0140353P.
23-JUN-1999; 99US-0140354P.
24-JUN-1999; 99US-0140695P.
28-JUN-1999; 99US-0140823P.
29-JUN-1999; 99US-0140991P.
30-JUN-1999; 99US-0141287P.
01-JUL-1999; 99US-0141842P.
01-JUL-1999; 99US-0142154P.
02-JUL-1999; 99US-0142055P.
06-JUL-1999; 99US-0142390P.
08-JUL-1999; 99US-0142803P.
09-JUL-1999; 99US-0142920P.
12-JUL-1999; 99US-0143977P.
13-JUL-1999; 99US-0143542P.
14-JUL-1999; 99US-0143624P.
15-JUL-1999; 99US-0144005P.
16-JUL-1999; 99US-0144085P.
16-JUL-1999; 99US-0144086P.
19-JUL-1999; 99US-0144325P.
19-JUL-1999; 99US-0144331P.
19-JUL-1999; 99US-0144332P.
19-JUL-1999; 99US-0144333P.
19-JUL-1999; 99US-0144334P.
19-JUL-1999; 99US-0144335P.
20-JUL-1999; 99US-0144352P.
20-JUL-1999; 99US-0144632P.
20-JUL-1999; 99US-0144884P.
21-JUL-1999; 99US-0144814P.
21-JUL-1999; 99US-0145086P.


```

XX SQ Sequence 42 AA;
Query Match 25.0%; Score 5; DB 5; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PRGAP 5
Db 14 PRGAP 18

RESULT 23
ADF94732
ID ADF94732 standard; protein; 42 AA.
XX AC ADF94732;
XX DT 26-FEB-2004 (first entry)
XX DE Hepatitis D virus type II antigen protein fragment, SEQ ID 7.
XX KW Virucide; hepatitis D virus antigen; HDV antigen; RNA polymerase II.
XX OS Hepatitis D virus.
XX PN WO200268655-A1.
XX PR 06-SEP-2002.
XX PF 27-FEB-2002; 2002WO-JP001817.
XX PR 27-FEB-2001; 2001JP-00053163.
XX PA (CIRC-) CIRCLE PROMOTION SCI & ENG.
XX PA (KYOW) KYOWA HAKKO KOGYO KK.
XX PI Handa H, Yamaguchi Y;
XX DR WPI; 2002-698673/75.
XX PT Non-infective cell-requiring method of searching for hepatitis D remedies
XX PT in a system containing e.g. polypeptide binding to RNA polymerase II with
XX PT measuring of binding level for indication.
XX PS Claim 5; SEQ ID NO 7; 84pp; Japanese.
XX CC The present invention relates to a method for searching hepatitis D
XX CC remedies. The method comprises adding a test compound to a system
XX CC containing e.g. a hepatitis D virus antigen (HDV antigen) and RNA
XX CC polymerase II, measuring the binding level of HDV antigen binding to RNA
XX CC polymerase II, and comparing with a control for selection of a compound
XX CC inhibiting the binding. The method is for screening remedies for
XX CC hepatitis D including for hepatitis virus D accompanying hepatitis B,
XX CC such as gene expression promoters applicable in regeneration medicine for
XX CC treating organs or tissues. The present sequence is a HDV antigen
XX CC sequence.
XX SQ Sequence 42 AA;
Query Match 25.0%; Score 5; DB 5; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PRGAP 5
Db 14 PRGAP 18

RESULT 24
AAW57222
ID AAW57222 standard; protein; 44 AA.
XX
```

```

AC AAW57222;
XX DT 04-AUG-1998 (first entry)
XX DE Targeting vector Tg-2/Not containing PGK-neomycin-resistant gene protein.
XX KW Targeting vector; Tg-2/Not; nucleobindin; screening; nephritis;
XX KW blood vessel; inflammation; PGK-neomycin-resistant gene.
XX OS Synthetic.
XX PN JP10117633-A.
XX PD 12-MAY-1998.
XX PF 21-OCT-1996; 96JP-00298219.
XX PR 21-OCT-1996; 96JP-00298219.
XX PA (MITS-) MITSUI SEIYAKU KOGYO KK.
XX PA (KANA/) KANAI Y.
XX PA (FUJI-) FUJITA GAKUEN.
XX DR WPI; 1998-325762/29.
XX DR N-PSDB; AAV28851.
XX PT Gene-deleted animal e.g. mouse - useful for screening therapeutic agents
XX PT for diseases such as blood vessel inflammation and nephritis.
XX PS Example 2; Fig 3; 11pp; Japanese.
XX CC The present sequence represents a protein from a targeting vector Tg-
XX CC 2/Not containing a PGK-neomycin-resistant gene, from the present
XX CC invention. The targeting vector is used in an example of the present
XX CC invention for producing a gene-deleted animal e.g. a mouse comprising no
XX CC gene to encode nucleobindin in which all or part of nucleobindin gene is
XX CC deleted or replaced by another gene. The mouse can be used for screening
XX CC therapeutic agents against diseases such as blood vessel inflammation and
XX CC nephritis.
XX SQ Sequence 44 AA;
Query Match 25.0%; Score 5; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PRGAP 5
Db 6 PRGAP 10

RESULT 25
AAV44951
ID AAV44951 standard; peptide; 46 AA.
XX AC AAV44951;
XX DT 23-MAY-2000 (first entry)
XX DE Xenopus thrombospondin 4 trimerisation domain.
XX KW Thrombospondin 4; TSP4; KDEL receptor inhibitor; heat shock protein;
XX KW immune response; oligomerisation domain; neoplasia; sarcoma; lymphoma;
XX KW leukaemia; melanoma; carcinoma; glioblastoma; astrocytoma; oncogene;
XX KW infectious disease; allergy; autoimmune disease.
XX OS Xenopus sp.
XX PN WO200006729-A1.
XX PD 10-FEB-2000.
XX PR 28-JUL-1999; 99WO-US017147.

```

XX PR 29-JUL-1998; 98US-00124671.
XX PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX PI Rothman JE, Mayhew M, Hoe MH;
XX XX WPI; 2000-195296/17.
XX DR
XX PT Inhibitors of the KDEL receptor which comprises an oligomerization domain
XX PT useful for promoting secretion of proteins which are normally retained
XX PT within the cell.
XX XX
XX PS Claim 12; Page 30; 87pp; English.
XX CC The patent discloses the use of KDEL receptor inhibitor to promote
XX CC secretion of proteins that are normally retained within the cell such as
XX CC heat shock proteins by inhibiting KDEL receptor-mediated return of
XX CC protein complexes to endoplasmic reticulum. This makes the secreted heat
XX CC shock proteins more accessible to the immune system and improves immune
XX CC response to a target antigen. The inhibitor protein comprises several
XX CC subunits where each subunit comprises an oligomerisation domain and has
XX CC at its carboxy terminus a region which binds to a KDEL receptor. The
XX CC target antigen may be associated with diseases including neoplasia such
XX CC as sarcoma, lymphoma, leukemia, melanoma, carcinoma, glioblastoma and
XX CC astrocytoma, with defective tumour suppressor genes, oncogenes,
XX CC infectious diseases, allergy or autoimmune diseases. The present sequence
XX CC is Xenopus thrombospondin 4 (TSP4) trimerisation domain. Oligomers formed
XX CC via oligomerisation domain of TSP4 are used to produce high avidity
XX CC binding protein which bind to KDEL receptor
XX XX
XX SQ Sequence 46 AA;
XX
XX Query Match 25.0%; Score 5; DB 3; Length 46;
XX Best Local Similarity 100.0%; Pred. No. 2.5e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 16 ETMFL 20
XX |||||
XX DB 31 ETMFL 35
XX
XX RESULT 26
XX AAM21142
XX ID AAM21142 standard; protein; 46 AA.
XX XX
XX AC AAM21142;
XX XX
XX DT 12-OCT-2001 (first entry)
XX XX
XX DE Peptide #7576 encoded by probe for measuring cervical gene expression.
XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX KW cervical cancer.
XX OS Homo sapiens.
XX XX
XX EN WO200157278-A2.
XX XX
XX PD 09-AUG-2001.
XX XX
XX PF 30-JAN-2001; 2001WO-US000670.
XX XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX XX WPI; 2001-488901/53.
XX DR
XX XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human cervical epithelial cells.
XX PS Claim 27; SEQ ID NO 20235; 487pp; English.
XX CC The present invention relates to human single exon nucleic acid probes
XX CC (SENP: see AAM10068-AA128459). The present sequence is a peptide encoded
XX CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
XX CC can be used to produce a single exon microarray, which can be used for
XX CC measuring human gene expression in a sample derived from human cervical
XX CC epithelial cells. By measuring gene expression, the probes are therefore
XX CC useful in grading and/or staging of diseases of the cervix, notably
XX CC cervical cancer. Note: The sequence data for this patent did not form
XX CC part of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX XX WPI; 2001-488901/53.
XX XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human cervical epithelial cells.
XX PS Claim 27; SEQ ID NO 25968; 487pp; English.
XX XX
XX CC The present invention relates to human single exon nucleic acid probes
XX CC (SENP: see AAM10068-AA128459). The present sequence is a peptide encoded
XX CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
XX CC can be used to produce a single exon microarray, which can be used for
XX CC measuring human gene expression in a sample derived from human cervical
XX CC epithelial cells. By measuring gene expression, the probes are therefore
XX CC useful in grading and/or staging of diseases of the cervix, notably
XX CC cervical cancer. Note: The sequence data for this patent did not form
XX CC part of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX XX
XX SQ Sequence 46 AA;
XX
XX Query Match 25.0%; Score 5; DB 4; Length 46;
XX Best Local Similarity 100.0%; Pred. No. 2.5e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 13 QMLET 17
XX |||||
XX DB 32 QMLET 36
XX
XX RESULT 27
XX AAM15409
XX ID AAM15409 standard; protein; 46 AA.
XX XX
XX AC AAM15409;
XX XX
XX DT 12-OCT-2001 (first entry)
XX XX
XX DE Peptide #1843 encoded by probe for measuring cervical gene expression.
XX XX Probe; human; microarray; gene expression; cervical epithelial cell;
XX KW cervical cancer.
XX OS Homo sapiens.
XX XX
XX EN WO200157278-A2.
XX XX
XX PD 09-AUG-2001.
XX XX
XX PF 30-JAN-2001; 2001WO-US000670.
XX XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX XX WPI; 2001-488901/53.
XX DR
XX XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human cervical epithelial cells.
XX PS Claim 27; SEQ ID NO 20235; 487pp; English.
XX CC The present invention relates to human single exon nucleic acid probes
XX CC (SENP: see AAM10068-AA128459). The present sequence is a peptide encoded
XX CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
XX CC can be used to produce a single exon microarray, which can be used for
XX CC measuring human gene expression in a sample derived from human cervical
XX CC epithelial cells. By measuring gene expression, the probes are therefore
XX CC useful in grading and/or staging of diseases of the cervix, notably
XX CC cervical cancer. Note: The sequence data for this patent did not form
XX CC part of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX XX

CC by one such probe. The SNPs are derived from human HeLa cells. The SNPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 46 AA;

Query Match 25.0%; Score 5; DB 4; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
 Db 20 RGAPM 24

RESULT 28
 ABB43459
 ID ABB43459 standard; peptide; 46 AA.
 XX
 AC ABB43459;
 XX
 DT 04-FEB-2002 (first entry)
 XX

DE Peptide #10965 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 KW Homo sapiens.
 XX
 OS WO200157277-A2.
 XX
 PN 09-AUG-2001.
 XX

PF 30-JAN-2001; 2001WO-US000669.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human fetal liver.

PS Claim 27; SEQ ID NO 36094; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 46 AA;

Query Match 25.0%; Score 5; DB 4; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
 Db 20 RGAPM 24

RESULT 30
 AAM27897
 ID AAM27897 standard; protein; 46 AA.
 XX
 AC AAM27897;
 XX
 DT 17-OCT-2001 (first entry)
 XX

DE Peptide #1920 encoded by human foetal liver single exon probe.
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 KW Homo sapiens.
 XX
 OS WO200157277-A2.
 XX
 PN 09-AUG-2001.
 XX

PF 30-JAN-2001; 2001WO-US000669.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human fetal liver.

PS Claim 27; SEQ ID NO 36094; 639pp + Sequence Listing; English.
 XX The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 46 AA;

Query Match 25.0%; Score 5; DB 4; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
 Db 32 QMLET 36

RESULT 29
 ABB34414
 ID ABB34414 standard; peptide; 46 AA.
 XX
 AC ABB34414;
 XX
 DT 04-FEB-2002 (first entry)
 XX

DE Peptide #1920 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 KW Homo sapiens.
 XX
 OS WO200157277-A2.
 XX
 PN 09-AUG-2001.
 XX

PF 30-JAN-2001; 2001WO-US000669.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human fetal liver.

PS Claim 27; SEQ ID NO 27049; 639pp + Sequence Listing; English.
 XX The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 46 AA;

Query Match 25.0%; Score 5; DB 4; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
 Db 20 RGAPM 24

RESULT 30
 AAM27897
 ID AAM27897 standard; protein; 46 AA.
 XX
 AC AAM27897;
 XX
 DT 17-OCT-2001 (first entry)
 XX

DE Peptide #1920 encoded by human foetal liver single exon probe.
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 KW Homo sapiens.
 XX
 OS WO200157277-A2.
 XX
 PN 09-AUG-2001.
 XX

PF 30-JAN-2001; 2001WO-US000669.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human fetal liver.

PS Claim 27; SEQ ID NO 27049; 639pp + Sequence Listing; English.
 XX The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 46 AA;

Query Match 25.0%; Score 5; DB 4; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
 Db 20 RGAPM 24

RESULT 30
 AAM27897
 ID AAM27897 standard; protein; 46 AA.
 XX
 AC AAM27897;
 XX
 DT 17-OCT-2001 (first entry)
 XX

DE Peptide #1920 encoded by human foetal liver single exon probe.
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 KW Homo sapiens.
 XX
 OS WO200157277-A2.
 XX
 PN 09-AUG-2001.
 XX

PF 30-JAN-2001; 2001WO-US000669.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human fetal liver.

PS Claim 27; SEQ ID NO 27049; 639pp + Sequence Listing; English.
 XX The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 46 AA;

Query Match 25.0%; Score 5; DB 4; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
 Db 20 RGAPM 24

RESULT 30
 AAM27897
 ID AAM27897 standard; protein; 46 AA.
 XX
 AC AAM27897;
 XX
 DT 17-OCT-2001 (first entry)
 XX

DE Peptide #1920 encoded by human foetal liver single exon probe.
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 KW Homo sapiens.
 XX
 OS WO200157277-A2.
 XX
 PN 09-AUG-2001.
 XX

PF 30-JAN-2001; 2001WO-US000669.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human fetal liver.

PS Claim 27; SEQ ID NO 27049; 639pp + Sequence Listing; English.
 XX The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 46 AA;

Query Match 25.0%; Score 5; DB 4; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE Peptide #1934 encoded by probe for measuring placental gene expression.
XX Probe; microarray; human; placenta; antenatal diagnosis;
KW Genetic disorder.
XX Homo sapiens.
XX OS
XX WO200157272-A2.
PN 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000663.
PF 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
XX Claim 27; SEQ ID NO 28166; 654pp; English.
XX The present invention relates to single exon nucleic acid probes (SENP:
CC see AA131315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
XX SQ Sequence 46 AA;
Query Match 25.0%; Score 5; DB 4; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RGAPM 6
Db 20 RGAPM 24
RESULT 31
AAM37347
ID AAM37347 standard; protein; 46 AA.
XX
XX AAM37347;
XX
XX 17-OCT-2001 (first entry)
XX
XX Peptide #11384 encoded by probe for measuring placental gene expression.
DE Probe; microarray; human; placenta; antenatal diagnosis;
KW Genetic disorder.
KW Homo sapiens.
XX
XX WO200157272-A2.
PN 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000663.
PF 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR

PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
XX Claim 27; SEQ ID NO 37616; 654pp; English.
XX The present invention relates to single exon nucleic acid probes (SENP:
CC see AA131315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
XX SQ Sequence 46 AA;
Query Match 25.0%; Score 5; DB 4; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 QMLET 17
Db 32 QMLET 36
RESULT 32
ABB29251
ID ABB29251 standard; peptide; 46 AA.
XX
XX ABB29251;
XX
XX 01-FEB-2002 (first entry)
XX Peptide #1902 encoded by breast cell single exon nucleic acid probe.
DE Human; microarray; single exon probe; gene expression; breast; disease;
XX cancer.
XX Homo sapiens.
XX OS
XX WO200157271-A2.
PN 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000662.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.

XX Claim 27; SEQ ID NO 12219; 327pp + Sequence Listing; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting the
 CC probes with a collection of detectably labelled nucleic acids derived
 CC from mRNA of human breast, and then measuring the label bound to each
 CC probe of the microarray. The probes are useful for verifying the
 CC expression of regions of genomic DNA predicted to encode proteins. They
 CC are useful for gene discovery, and for determining predisposition and/or
 CC prognosing breast disease. Gene expression analysis is useful for
 CC assessing the toxicity of chemical agents on cells. The microarray of
 CC this invention presents a far greater diversity of probes for measuring
 CC gene expression, with far less bias than expressed sequence tag
 CC microarrays. The method is suitable for rapid production of functional
 CC information from genomic sequence. The present sequence is a peptide
 CC encoded by a single exon nucleic acid probe of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 46 AA;
 SQ

Query Match 25.0%; Score 5; DB 4; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
 Db |||||
 20 RGAPM 24

RESULT 33
 ABB19826
 ID ABB19826 standard; protein; 46 AA.
 XX ABB19826;
 AC
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Protein #1825 encoded by probe for measuring heart cell gene expression.
 XX Human; gene expression; heart; microarray; vascular system;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease.
 XX Homo sapiens.
 OS
 XX WO200157274-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US000666.
 PF
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488999/53.
 PI Single exon nucleic acid probes for analyzing gene expression in human
 XX hearts.
 XX Claim 15; SEQ ID NO 21596; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 46 AA;
 SQ

Query Match 25.0%; Score 5; DB 4; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
 Db |||||
 20 RGAPM 24

RESULT 34
 AAM67599
 ID AAM67599 standard; protein; 46 AA.
 XX AAM67599;
 AC
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27905.
 XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX Homo sapiens.
 OS
 XX WO200157276-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US000668.
 PF
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488900/53.
 DR Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human bone marrow.
 XX Example 4; SEQ ID NO 27905; 658pp + Sequence Listing; English.
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention
 XX Sequence 46 AA;
 XX

KW hypercholesterolaemia; coronary heart disease.
 XX Homo sapiens.
 OS
 PN WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000664.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 21-AUG-2000; 2000US-00632366.
 PR 03-AUG-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488898/53.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human adult liver.
 XX
 PS Claim 27; SEQ ID NO 27894; 658pp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult liver.
 CC (I) may be used for predicting, measuring and displaying gene expression
 CC in samples derived from human adult liver. The genes identified may be
 CC involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 CC associated with coronary heart disease. ABG47348-ABG59930 represent human
 CC liver single exon encoded peptides of the invention. Note: The sequence
 CC information for this patent does not appear in the printed specification
 CC but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 46 AA;
 Query Match 25.0%; Score 5; DB 4; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RGAPM 6
 Db |||||
 20 RGAPM 24
 RESULT 38
 AAM03170
 ID AAM03170 standard; protein; 46 AA.
 AC
 XX AAM03170;
 XX
 DT 09-OCT-2001 (first entry)
 XX
 DE Peptide #1852 encoded by probe for measuring breast gene expression.
 XX Probe; human; breast disease; breast cancer; development disorder;
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
 KW
 XX Homo sapiens.
 OS
 XX
 PN WO200157270-A2.
 XX
 PD 09-AUG-2001.

XX
 PF 29-JAN-2001; 2001WO-US000661.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-476286/51.
 DR
 XX Novel single exon nucleic acid probe used to measuring gene expression in
 PT a human breast.
 XX
 PS Claim 27; SEQ ID NO 11910; 322pp; English.
 XX
 CC The present invention relates to novel single exon nucleic acid probes
 CC (see AA100010-AA110067). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for measuring human gene expression in
 CC a human breast sample, where the probe hybridises at high stringency to a
 CC nucleic acid expressed in the human breast. The probes are useful for
 CC predicting, diagnosing, grading, staging, monitoring and prognosing
 CC diseases of the human breast, particularly those diseases with polygenic
 CC aetiology. The diseases include: breast cancer, disorders of development,
 CC inflammatory diseases of the breast, fibrocystic changes, proliferative
 CC breast disease and non-carcinoma tumours. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 46 AA;
 Query Match 25.0%; Score 5; DB 4; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RGAPM 6
 Db |||||
 20 RGAPM 24
 RESULT 39
 ABG46222
 ID ABG46222 standard; peptide; 46 AA.
 XX
 AC ABG46222;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 35987.
 XX
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; pulmonary syndrome;
 KW primary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.

PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 27; SEQ ID NO 35887; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 46 AA;
XX
XX Query Match 25.0%; Score 5; DB 5; Length 46;
XX Best Local Similarity 100.0%; Pred. No. 2.5e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 13 QMLET 17
XX Db 32 QMLET 36
XX
XX RESULT 40
XX .ABG37192

ID ABG37192 standard; peptide; 46 AA.
XX
AC ABG37192;
XX
XX 19-AUG-2002 (first entry)
DT
DE Human peptide encoded by genome-derived single exon probe SEQ ID 26857.
XX
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
XX Homo sapiens.
OS
XX WO200186003-A2.
PN
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US000665.
PF
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
PR
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
PT
XX Claim 27; SEQ ID NO 26857; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 46 AA;
XX
XX Query Match 25.0%; Score 5; DB 5; Length 46;
XX Best Local Similarity 100.0%; Pred. No. 2.5e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 13 QMLET 17
XX Db 32 QMLET 36
XX
XX RESULT 40
XX .ABG37192

CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagenet syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a peptide/protein encoded by a single exon probe of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 46 AA;

Query Match 25.0%; Score 5; DB 5; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
 DB 20 RGAPM 24
 |||||

RESULT 41
 ABB38152
 ID ABB38152 standard; peptide; 47 AA.

XX AC ABB38152;
 XX AC
 DT 04-FEB-2002 (first entry)

XX Peptide #5658 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000669.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human fetal liver.

XX Claim 27; SEQ ID NO 30787; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring

XX human gene expression in a sample derived from human foetal liver. The

XX single exon nucleic acid probes may be used for predicting, measuring and

XX displaying gene expression in samples derived from human fetal liver. The

XX present sequence is a peptide encoded by a single exon nucleic acid probe

XX of the invention. Note: The sequence data for this patent did not form

Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 DB 13 PRGAP 17
 |||||

RESULT 42

AAAM31580
 ID AAAM31580 standard; protein; 47 AA.

XX AC AAAM31580;

XX 17-OCT-2001 (first entry)

XX Peptide #5617 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;

XX Genetic disorder.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000663.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human placenta.

XX Claim 27; SEQ ID NO 31849; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs;

XX see AAI31315-AAI57546). The present sequence is a peptide encoded by one

XX such probe. The probes are useful for producing a microarray for

XX predicting, measuring and displaying gene expression in samples derived

XX from human placenta. The probes are useful for antenatal diagnosis of

XX human genetic disorders

XX Sequence 47 AA;

Query Match 25.0%; Score 5; DB 4; Length 47;

Best Local Similarity 100.0%; Pred. No. 2.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 DB 13 PRGAP 17
 |||||

RESULT 43

ABB23358

ID ABB23358 standard; protein; 47 AA.

XX AC ABB23358;

XX 23-JAN-2002 (first entry)

```
XX Protein #5357 encoded by probe for measuring heart cell gene expression.
DE Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
OS Homo sapiens.
XX WO200157274-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000666.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488999/53.
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts.
XX Claim 15; SEQ ID NO 25128; 530pp; English.
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX ABA21535-ABA41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease.
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 47 AA;
XX
XX Query Match 25.0%; Score 5; DB 4; Length 47;
XX Best Local Similarity 100.0%; Pred.No. 2.6e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 PRGAP 5
XX Db 13 PRGAP 17
XX
XX RESULT 44
XX AAW71301
XX ID AAW71301 standard; protein; 47 AA.
XX AC AAW71301;
XX
XX DT 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 31607.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX Homo sapiens.
XX
```

```
PN WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000668.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488999/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 31607; 659pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is a
XX protein encoded by one of the probes of the invention
XX
XX Sequence 47 AA;
XX
XX Query Match 25.0%; Score 5; DB 4; Length 47;
XX Best Local Similarity 100.0%; Pred.No. 2.6e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 PRGAP 5
XX Db 13 PRGAP 17
XX
XX RESULT 45
XX AAM58788
XX ID AAM58788 standard; protein; 47 AA.
XX
XX AC AAM58788;
XX
XX DT 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 30893.
XX
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
```

XX Penn SG, Hanzel DX, Chen W, Rank DR;
 XX WPI; 2001-483446/52.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.
 XX Example 4; SEQ ID NO 30893; 650pp + Sequence Listing; English.
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention
 XX
 XX Sequence 47 AA;
 SQ
 Query Match 25.0%; Score 5; DB 4; Length 47;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PRGAP 5
 Db 13 PRGAP 17
 RESULT 46
 ABG53012
 ID ABG53012 standard; peptide; 47 AA.
 AC ABG53012;
 XX 25-FEB-2003 (first entry)
 DT Human liver peptide, SEQ ID NO 31660.
 XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX Homo sapiens.
 OS
 XX WO200157273-A2.
 PN 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US000664.
 PF 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DX, Chen W, Rank DR;
 PI WPI; 2001-48898/53.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human adult liver.
 PT
 XX Claim 27; SEQ ID NO 31660; 658pp; English.
 XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high

CC stringency to a nucleic acid molecule expressed in the human adult liver.
 CC (I) may be used for predicting, measuring and displaying gene expression
 CC in samples derived from human adult liver. The genes identified may be
 CC involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 CC associated with coronary heart disease. ABG47348-ABG59930 represent human
 CC liver single exon encoded peptides of the invention. Note: The sequence
 CC information for this patent does not appear in the printed specification
 CC but was obtained in electronic format directly from WIFO at
 CC ftp.wifo.int/pub/published_pct_sequences
 XX
 XX Sequence 47 AA;
 SQ
 Query Match 25.0%; Score 5; DB 4; Length 47;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PRGAP 5
 Db 13 PRGAP 17
 RESULT 47
 ABG41101
 ID ABG41101 standard; peptide; 47 AA.
 XX AC ABG41101;
 XX 19-AUG-2002 (first entry)
 DT Human peptide encoded by genome-derived single exon probe SEQ ID 30766.
 XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
 DE chronic obstructive pulmonary disease; interstitial lung disease;
 XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX Homo sapiens.
 OS
 XX WO200186003-A2.
 PN 15-NOV-2001.
 PD
 XX 30-JAN-2001; 2001WO-US000665.
 PF 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2002-114183/15.
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 PT
 XX Claim 27; SEQ ID NO 30766; 634pp; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their

complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridise at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberculous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and bialine membrane disease. The present sequence is a peptide/protein encoded by a single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct_sequences

XX	WPI; 2001-639362/73.
DR	N-PSDB; AAS83588.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
XX	Claim 20; SEQ ID NO 49760; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC	amino acid sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 50 AA;
SQ	
Query Match	25.0%; Score 5; DB 4; Length 50;
Best Local Similarity	100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Oy	1 PRGAP 5
Db	16 PRGAP 20
RESULT 49	
AAG57910	
ID	AAG57910 standard; protein; 54 AA.
XX	
AC	AAG57910;
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 74690.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
OS	
OS	Arabidopsis thaliana.
PX	
PN	EPI033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 24-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135629P.
PR 27-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136392P.
PR 01-JUN-1999; 99US-0136782P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151338P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.

QY	DB	Query Match	Best Local Similarity	25.0%; Score 5; DB 3; Length 54;	Matches	5; Conservative	100.0%; Pred. No. 2.9e+02;	Mismatches	0;	Indels	0;	Gaps	0;
PR	13-OCT-1999;	99US-0159294P.											
PR	13-OCT-1999;	99US-0159295P.											
PR	14-OCT-1999;	99US-0159329P.											
PR	14-OCT-1999;	99US-0159330P.											
PR	14-OCT-1999;	99US-0159331P.											
PR	14-OCT-1999;	99US-0159637P.											
PR	14-OCT-1999;	99US-0159638P.											
PR	18-OCT-1999;	99US-0159584P.											
PR	21-OCT-1999;	99US-0160741P.											
PR	21-OCT-1999;	99US-0160769P.											
PR	21-OCT-1999;	99US-0160770P.											
PR	21-OCT-1999;	99US-0160814P.											
PR	21-OCT-1999;	99US-0160815P.											
PR	22-OCT-1999;	99US-0160980P.											
PR	22-OCT-1999;	99US-0160981P.											
PR	25-OCT-1999;	99US-0161404P.											
PR	25-OCT-1999;	99US-0161405P.											
PR	25-OCT-1999;	99US-0161406P.											
PR	26-OCT-1999;	99US-0161359P.											
PR	26-OCT-1999;	99US-0161360P.											
PR	26-OCT-1999;	99US-0161361P.											
PR	28-OCT-1999;	99US-0161920P.											
PR	28-OCT-1999;	99US-0161992P.											
PR	28-OCT-1999;	99US-0161993P.											
PR	29-OCT-1999;	99US-0162142P.											
PR	29-OCT-1999;	99US-0162143P.											
PR	29-OCT-1999;	99US-0162144P.											
PR	29-OCT-1999;	99US-0162145P.											
PR	29-OCT-1999;	99US-0162146P.											
PR	29-OCT-1999;	99US-0162147P.											
PR	29-OCT-1999;	99US-0162148P.											
PR	29-OCT-1999;	99US-0162149P.											
PR	29-OCT-1999;	99US-0162150P.											
PR	29-OCT-1999;	99US-0162151P.											
PR	29-OCT-1999;	99US-0162152P.											
PR	29-OCT-1999;	99US-0162153P.											
PR	29-OCT-1999;	99US-0162154P.											
PR	29-OCT-1999;	99US-0162155P.											
PR	29-OCT-1999;	99US-0162156P.											
PR	29-OCT-1999;	99US-0162157P.											
PR	29-OCT-1999;	99US-0162158P.											
PR	29-OCT-1999;	99US-0162159P.											
PR	29-OCT-1999;	99US-0162160P.											
PR	29-OCT-1999;	99US-0162161P.											
PR	29-OCT-1999;	99US-0162162P.											
PR	29-OCT-1999;	99US-016											

```
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 22-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145913P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151067P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 28-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 25.0%; Score 5; DB 3; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
Db 47 QMLET 51

RESULT 51
ABUG2915
ID ABUG2915 standard; protein; 54 AA.
XX
AC ABUG2915;
XX
DT 16-SEP-2003 (first entry)
XX
DE Mouse osteoblast differentiation promoting factor (OBDFP) fragment.
XX
KW Mouse; osteoblast differentiation promoting factor; OBDFP;
KW osteoblast differentiation; cell morphology; esterase expression;
KW bone remodeling; osteoporosis; osteomalacia; hypercalcaemia.
XX
OS Mus musculus.
XX
PN US2003044825-A1.
XX
PD 06-MAR-2003.
XX
PF 22-MAY-2002; 2002US-00152031.
XX
PR 23-MAY-2000; 2000JP-00139467.
PR 22-MAY-2001; 2001US-0292318P.
XX
(PA (IMAI/) IMAI Y.
PA (AKAT/) AKATSUKA H.
PA (KAWA/) KAWAI E.
PA (OMOR/) OMORI K.
PA (YANA/) YANAKA N.
PA (SAKU/) SAKURAI N.
XX
PI Imai Y, Akatsuka H, Kawai E, Omori K, Yanaka N, Sakurai N;
XX WPI; 2003-521747/49.
XX
PT New bone metabolism-related proteins and nucleic acids encoding the
PT proteins, useful in the activation process of the osteoblasts and
PT induction of successive recruitment of the osteoclasts in bone remodeling
PT process.
XX
PS Example 7; Fig 6; 31pp; English.
XX
```

CC The invention describes an isolated murine or human polypeptide (I). The
 CC nucleic acid is useful for suppressing or strengthening the promotion of
 CC differentiation of osteoblast, induction of morphological change of a
 CC cell, or expression of an esterase activity. The antibody is useful for
 CC neutralising a function or an activity of the polypeptide. Detecting a
 CC function or an activity of the polypeptide or the nucleic acid is useful
 CC for screening or identifying compounds, e.g. pharmaceutical compounds,
 CC which modulate the function or activity of the polypeptide. The
 CC identified compounds are valuable for selling. The polypeptides are also
 CC useful in the activation process of the osteoblasts and induction of
 CC successive recruitment of the osteoclasts in bone remodeling process. The
 CC polypeptides and nucleic acids are useful in studies on pathological
 CC states, diagnostics, therapeutic and prophylactic treatment, and research
 CC and development of pharmaceuticals for diseases such as osteoporosis,
 CC osteomalacia or hypercalcaemia. This is the amino acid sequence of mouse
 CC osteoblast differentiation promoting factor (OBDFP) fragment used in
 CC comparison with known enzymes during the functional analysis of OBDFP
 XX
 XX
 SQ Sequence 54 AA;

Query Match 25.0%; Score 5; DB 6; Length 54;
 Best Local Similarity 100.0%; Pred.No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGAPM 6
 |||||
 Db 4 RGAPM 8

RESULT 52

AAG60769

ID AAG60769 standard; protein; 55 AA.

XX AAG60769;

AC AAG60769;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 78752.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.
 XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.
 XX 25-FEB-1999; 99US-0121825P.
 XX 05-MAR-1999; 99US-0123180P.
 XX 09-MAR-1999; 99US-0123548P.
 XX 23-MAR-1999; 99US-0125788P.
 XX 25-MAR-1999; 99US-0126264P.
 XX 29-MAR-1999; 99US-0126785P.
 XX 01-APR-1999; 99US-0127462P.
 XX 06-APR-1999; 99US-0128234P.
 XX 08-APR-1999; 99US-0128714P.
 XX 16-APR-1999; 99US-0129845P.
 XX 19-APR-1999; 99US-0130077P.
 XX 21-APR-1999; 99US-0130449P.
 XX 23-APR-1999; 99US-0130510P.
 XX 28-APR-1999; 99US-0130891P.
 XX 30-APR-1999; 99US-0131449P.
 XX 30-APR-1999; 99US-0132048P.
 XX 04-MAY-1999; 99US-0132407P.
 XX 05-MAY-1999; 99US-0132484P.
 XX 06-MAY-1999; 99US-0132485P.
 XX 06-MAY-1999; 99US-0132486P.
 XX 07-MAY-1999; 99US-0132487P.
 XX 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136352P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142392P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.


```
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 27-JUL-1999; 99US-0145319P.
PR 28-JUL-1999; 99US-0145351P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149368P.
PR 18-AUG-1999; 99US-0149175P.
PR 20-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157863P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.

PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 25.0%; Score 5; DB 3; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.9e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
    |||||
Db 48 QMLET 52

RESULT 53
ABP64607
ID ABP64607 standard; protein; 57 AA.
XX
AC ABP64607;
XX
DT 04-NOV-2002 (first entry)
XX
DE Human ORF977.
XX
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnery;
KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder.
XX
OS Homo sapiens.
XX
US2002082206-A1.
XX
PD 27-JUN-2002.
XX
PF 30-MAY-2001; 2001US-00867550.
XX
PR 30-MAY-2000; 2000US-0208427P.
XX
PA (LEAC/) LEACH M D.
PA (MEHR/) MEHRABAN F.
PA (CONL/) CONLEY P B.
PA (TOPP/) TOPPER J N.
PA (LAWD/) LAW D.
XX
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
WPI: 2002-626554/67.
XX
N-PSDB; ABQ99170.
XX
PT New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.
XX
PS Claim 10; SEQ ID NO 1954; 78pp; English.
XX
CC The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
```

CC condition associated with an OREFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/sequence.html?docID=20020082206
XX
XX
SQ Sequence 57 AA;

Query Match 25.0%; Score 5; DB 5; Length 57;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 VCOML 15
| | | | |
DB 22 VCOML 26

RESULT 54
AAU46690
ID AAU46690 standard; protein; 62 AA.
XX
AC AAU46690;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #7586.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI: 2001-616774/71.
DR N-PSDB; AAS59534.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 7885; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and

CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 62 AA;

Query Match 25.0%; Score 5; DB 4; Length 62;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRGAP 5
| | | | |
DB 8 PRGAP 12

RESULT 55
ABM43209
ID ABM43209 standard; protein; 62 AA.
XX
AC ABM43209;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #7885.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglas J;
XX
XX WPI: 2003-381789/36.
XX N-PSDB; ACF64463.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX or for stimulating an immune response specific for a P. acnes protein.
XX Example 1; SEQ ID NO 7885; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne

CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 62 AA;

Query Match 25.0%; Score 5; DB 6; Length 62;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5

Db 8 PRGAP 12

RESULT 56

AAG60768

ID AAG60768 standard; protein; 63 AA.

XX
AC AAG60768;

XX
DT 18-OCT-2000 (first entry)

XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 78751.

XX
KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX
OS Arabidopsis thaliana.

XX
PN EP1033405-A2.

XX
PD 06-SEP-2000.

XX
PF 25-FEB-2000; 2000EP-00301439.

XX
PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125789P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 18-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140153P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140693P.

PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140991P.

PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.

PR 02-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.

PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0142977P.

PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.

PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 20-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144884P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 21-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145132P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145951P.

Tue Oct 26 08:02:19 2004

PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147302P.
PR 04-AUG-1999; 99US-0147303P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155133P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 29-SEP-1999; 99US-0157117P.
PR 04-OCT-1999; 99US-0157753P.
PR 05-OCT-1999; 99US-0157863P.
PR 06-OCT-1999; 99US-0158029P.
PR 07-OCT-1999; 99US-0158232P.
PR 08-OCT-1999; 99US-0158369P.
PR 12-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.

PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 25.0%; Score 5; DB 3; Length 63;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 QMLET 17
Db 56 QMLET 60

RESULT 57
ABO58509
ID ABO58509 standard; protein; 63 AA.
XX AC ABO58509;
XX DT 29-JUL-2004 (first entry)
XX Human genome derived single exon protein #4743.
DE Human; gene expression; single exon probe; microarray;
XX KW alternative splicing event; genomic alteration.
KW OS Homo sapiens.
XX US2003194704-A1.
PD 16-OCT-2003.
PP 03-APR-2002; 2002US-00029386.
PP 03-APR-2002; 2002US-00029386.
PR (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX Penn SG, Rank DR, Hanzel DK;
WPI; 2004-119264/12.

New human genome-derived single exon nucleic acid probes useful for human gene expression analysis, for identifying or characterizing alternative splicing events, for assessing genomic alterations or as tools for surveying tissues.

Claim 45; SEQ ID NO 32143; 80pp; English.

The invention relates to a nucleic acid probe for measuring human gene expression, comprising any of the 27,400 fully defined nucleotide sequences in the specification, or their complements or fragments, and encoding at least 8 amino acids of any of the 688 amino acid sequences fully defined in the specification. The probe is a single exon probe that hybridizes under high stringency conditions to a nucleic acid molecule expressed in human cells or tissues. Also included are a spatially-addressable set of single exon nucleic acid probes for measuring human gene expression (comprising a plurality of single exon nucleic acid probes cited above, where each of the plurality of probes is separately and addressably isolatable or amplifiable from the plurality), a single exon microarray for measuring human gene expression, a method of measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above, methods of selling and/or licensing single exon probes or microarrays to a customer desiring to measure gene expression, a method of providing human gene expression data by subscription, and a computer-readable

CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterising
 CC alternative splicing events, in detecting and characterising gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe protein of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704
 CC
 XX SQ Sequence 63 AA;
 CC
 CC Query Match 25.0%; Score 5; DB 8; Length 63;
 CC Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 QY 1 PRGAP 5
 Db 15 PRGAP 19
 CC
 CC RESULT 58
 CC ABP00767
 CC ID ABP00767 standard; protein; 64 AA.
 CC XX AC ABP00767;
 CC XX
 CC XX 25-JUN-2002 (first entry)
 CC DE Human ORFX protein sequence SEQ ID NO:1516.
 CC XX
 CC KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 CC KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 CC KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 CC KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 CC KW hypertension; hypothyroidism; cholesterol ester storage disease;
 CC KW immune deficiency; immune disorder; infectious disease;
 CC KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 CC KW myasthenia gravis.
 CC XX OS Homo sapiens.
 CC XX WO200192523-A2.
 CC XX PD 06-DEC-2001.
 CC XX 29-MAY-2001; 2001WO-US010836.
 CC XX 30-MAY-2000; 2000US-0206132P.
 CC ER 29-AUG-2000; 2000US-0228716P.
 CC XX (CURA-) CURAGEN CORP.
 CC XX PA
 CC XX PI Shimkets RA, Leach MD;
 CC XX WPI; 2002-106308/14.
 CC DR N-PSDB; ABN16519.
 CC XX
 CC PT Novel human polypeptides and polynucleotides useful for diagnosing,
 CC PT preventing and treating cardiovascular disease, neurodegenerative,
 CC PT hyperproliferative disorders and autoimmune disorders.
 CC XX Disclosure; SEQ ID NO 1516; 1037pp; English.
 CC PS
 CC XX The present invention describes substantially purified human proteins
 CC CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX

CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 XX SQ Sequence 64 AA;
 CC
 CC Query Match 25.0%; Score 5; DB 5; Length 64;
 CC Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 QY 1 PRGAP 5
 Db 19 PRGAP 23
 CC
 CC RESULT 59
 CC AAG61788
 CC ID AAG61788 standard; protein; 65 AA.
 CC XX AC AAG61788;
 CC XX 18-OCT-2000 (first entry)
 CC DE Arabidopsis thaliana protein fragment SEQ ID NO: 80257.
 CC KW Protein identification; signal transduction pathway; metabolic pathway;
 CC KW hybridisation assay; Genetic mapping; Gene expression control; promoter;
 CC KW termination sequence.
 CC XX Arabidopsis thaliana.
 CC XX EP1033405-A2.
 CC PD 06-SEP-2000.
 CC XX 25-FEB-2000; 2000EP-00301439.
 CC XX 25-FEB-1999; 99US-0121825P.
 CC ER 05-MAR-1999; 99US-0123180P.
 CC ER 09-MAR-1999; 99US-0123548P.
 CC ER 23-MAR-1999; 99US-0125788P.
 CC ER 25-MAR-1999; 99US-0126264P.
 CC ER 29-MAR-1999; 99US-0126785P.
 CC ER 01-APR-1999; 99US-0127462P.
 CC ER 06-APR-1999; 99US-0128234P.
 CC ER 16-APR-1999; 99US-0128714P.
 CC ER 19-APR-1999; 99US-0129845P.
 CC ER 21-APR-1999; 99US-0130077P.
 CC ER 23-APR-1999; 99US-0130449P.
 CC ER 23-APR-1999; 99US-0130510P.
 CC ER 28-APR-1999; 99US-0130891P.
 CC ER 30-APR-1999; 99US-0131449P.
 CC ER 30-APR-1999; 99US-0132048P.
 CC ER 30-APR-1999; 99US-0132407P.
 CC ER 04-MAY-1999; 99US-0132484P.

Tue Oct 26 08:02:19 2004

us-10-066-965a-3.oligo.rag

```
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136352P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138084P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139482P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0132485P.
PR 22-JUL-1999; 99US-0132486P.
PR 22-JUL-1999; 99US-0132487P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 22-SEP-1999; 99US-0154779P.
PR 23-SEP-1999; 99US-0155139P.
PR 24-SEP-1999; 99US-0155486P.
PR 28-SEP-1999; 99US-0155659P.
PR 29-SEP-1999; 99US-0156458P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
```

PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 23-OCT-1999; 99US-0161404P.
 PR 23-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161520P.
 PR 28-OCT-1999; 99US-0161592P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 25.0%; Score 5; DB 3; Length 65;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
 Db 58 QMLET 62

RESULT 60
 ADF94731
 ID ADF94731 standard; protein; 66 AA.

AC ADF94731;

XX 26-FEB-2004 (first entry)

DT Hepatitis D virus type II antigen protein fragment, SEQ ID 6.

DE Virucide; hepatitis D virus antigen; HDV antigen; RNA polymerase II.

KW Hepatitis D virus.

OS WO200268655-A1.

PN 06-SEP-2002.

PD 27-FEB-2002; 2002WO-JP001817.

PF 27-FEB-2001; 2001JP-00053163.

PR (CIRC-) CIRCLE PROMOTION SCI & ENG.

PA (KYOW) KYOWA HAKKO KOGYO KK.

XX Handa H, Yamaguchi Y;

PI WPI; 2002-698673/75.

DR Non-infective cell-requiring method of searching for hepatitis D remedies
 in a system containing e.g. polypeptide binding to RNA polymerase II with
 measuring of binding level for indication.

PS Claim 5; SEQ ID NO 6; 84pp; Japanese.

XX The present invention relates to a method for searching hepatitis D
 remedies. The method comprises adding a test compound to a system
 containing e.g. a hepatitis D virus antigen (HDV antigen) and RNA
 polymerase II, measuring the binding level of HDV antigen binding to RNA
 polymerase II, and comparing with a control for selection of a compound
 inhibiting the binding. The method is for screening remedies for
 hepatitis D including for hepatitis virus D accompanying hepatitis B,
 such as gene expression promoters applicable in regeneration medicine for
 treating organs or tissues. The present sequence is a HDV antigen
 sequence.

XX Sequence 66 AA;

Query Match 25.0%; Score 5; DB 5; Length 66;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 Db 30 PRGAP 34

RESULT 61
 ADF94733
 ID ADF94733 standard; protein; 66 AA.

AC ADF94733;

XX 26-FEB-2004 (first entry)

DT Hepatitis D virus type III antigen protein fragment, SEQ ID 8.

DE Virucide; hepatitis D virus antigen; HDV antigen; RNA polymerase II.

KW Hepatitis D virus.

OS WO200268655-A1.

PN 06-SEP-2002.

PD 27-FEB-2002; 2002WO-JP001817.

PF 27-FEB-2001; 2001JP-00053163.

PR (CIRC-) CIRCLE PROMOTION SCI & ENG.

PA (KYOW) KYOWA HAKKO KOGYO KK.

XX Handa H, Yamaguchi Y;

PI WPI; 2002-698673/75.

DR Non-infective cell-requiring method of searching for hepatitis D remedies
 in a system containing e.g. polypeptide binding to RNA polymerase II with
 measuring of binding level for indication.

PS Claim 5; SEQ ID NO 8; 84pp; Japanese.

XX The present invention relates to a method for searching hepatitis D
 remedies. The method comprises adding a test compound to a system
 containing e.g. a hepatitis D virus antigen (HDV antigen) and RNA
 polymerase II, measuring the binding level of HDV antigen binding to RNA
 polymerase II, and comparing with a control for selection of a compound
 inhibiting the binding. The method is for screening remedies for
 hepatitis D including for hepatitis virus D accompanying hepatitis B,
 such as gene expression promoters applicable in regeneration medicine for
 treating organs or tissues. The present sequence is a HDV antigen
 sequence.

SQ Sequence 66 AA;

Query Match 25.0%; Score 5; DB 5; Length 66;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 Db 30 PRGAP 34

RESULT 62
 AAM83752
 ID AAM83752 standard; protein; 67 AA.

XX AAM83752;

DT 07-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen SEQ ID NO:11345.
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis.
KW Homo sapiens.
XX WO200157182-A2.
XX 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US0011354.
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-01981123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 23-AUG-2000; 2000US-0227182P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0228287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 21-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241878P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-483426/52.
 DR N-PSDB; AAK56533.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Claim 11; SEQ ID NO 11345; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I). Additionally, (I)
 CC nucleic acids may be used to culture the secreted (I), by inserting the
 CC protein. (I) proteins and polynucleotides may be used to express the
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 XX
 SQ Sequence 67 AA;
 Query Match 25.0%; Score 5; DB 4; Length 67;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PRGAP 5
 Db 36 PRGAP 40
 RESULT 63
 ABP64418
 ID ABP64418 standard; protein; 67 AA.
 XX
 AC ABP64418;
 XX
 DT 04-NOV-2002 (first entry)
 XX
 DE Human ORF788.
 XX
 KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnary;
 KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
 KW cancer; cardiovascular disease; allergy; autoimmune disease;
 KW wound healing; blood coagulation disorder; inflammatory disorder.
 XX
 OS Homo sapiens.
 XX
 PN US2002082206-A1.
 XX
 PD 27-JUN-2002.
 XX
 PF 30-MAY-2001; 2001US-00867550.
 XX
 PR 30-MAY-2000; 2000US-0208427P.
 XX

PA (LEAC/) LEACH M D.
 PA (MEHR/) MEHRABAN F.
 PA (CONL/) CONLEY P B.
 PA (TOPP/) TOPPER J N.
 PA (LAWD/) LAW D.
 XX
 PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
 XX
 DR WPI; 2002-626554/67.
 DR N-PSDB; ABQ98981.
 XX
 PT New polypeptide designated ORFX are present in human atherogenic cells
 PT and are useful to prevent and treat ORFX-associated disorders including
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or
 PT inflammatory disease.
 XX
 XX Claim 10; SEQ ID NO 1576; 78pp; English.
 XX
 CC The present invention relates to novel human ORFX polypeptides and their
 CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
 CC were discovered in human atherogenic cells, in particular in platelets
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
 CC many other tissues as well. Atherogenic cells are cells which have the
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
 CC nucleic acids are useful for treating or preventing a pathological
 CC condition associated with an ORFX-associated disorder, e.g. cancer,
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
 CC coagulation disorders or inflammatory disorders. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from the USPTO web site at
 CC seqdata.uspto.gov/sequence.html?DocID=20020082206
 XX
 SQ Sequence 67 AA;
 Query Match 25.0%; Score 5; DB 5; Length 67;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PRGAP 5
 Db 40 PRGAP 44
 RESULT 64
 AAU61722
 ID AAU61722 standard; protein; 68 AA.
 XX
 AC AAU61722;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #22618.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 XX 20-APR-2001; 2001WO-US012865.
 PF
 XX 21-APR-2000; 2000US-0199047P.
 PR
 PR 02-JUN-2000; 2000US-0208841P.
 PR
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI

```

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59620.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 22917; 1069pp; English.
PS
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
PS polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 68 AA;
SQ
Query Match 25.0%; Score 5; DB 4; Length 68;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
DB 61 PRGAP 65

RESULT 65
ABM58241
ID ABM58241 standard; protein; 68 AA.
XX
AC ABM58241;
XX
XX 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #22917.
XX
XX Acne vulgaris; antisporrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
XX WO2003033515-A1.
XX
XX 24-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032727.
XX
XX 15-OCT-2001; 2001US-00978825.
XX
XX (CORI-) CORIXA CORP.
XX
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Carter D;
PI Barth B, Vallieue-Douglass J;
XX
XX WPI; 2003-381789/36.
DR

```

```

DR N-PSDB; ACF64549.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
XX Example 1; SEQ ID NO 22917; 1481pp; English.
PS
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 68 AA;
SQ
Query Match 25.0%; Score 5; DB 6; Length 68;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
DB 61 PRGAP 65

RESULT 66
AAG57909
ID AAG57909 standard; protein; 70 AA.
XX
AC AAG57909;
XX
XX 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 74689.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR

```

PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0136229P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138340P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140821P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 27-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154018P.
PR 20-SEP-1999; 99US-0154039P.
PR 22-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.

```

PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159233P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 25.0%; Score 5; DB 3; Length 70;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
DB 63 QMLET 67

RESULT 67
AAO10516
ID AAO10516 standard; protein; 70 AA.
AC AAO10516;
XX
XX 06-NOV-2001 (first entry)
DE Human polypeptide SEQ ID NO 24408.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
XX (HYSB-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-514838/56.
DR
XX N-PSDB; AAI90447.
XX

PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 20; SEQ ID NO 24408; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 70 AA;
Query Match 25.0%; Score 5; DB 4; Length 70;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
DB 1 PRGAP 5

RESULT 68
AAU64459
ID AAU64459 standard; protein; 71 AA.
XX
XX AAU64459;
XX
XX 27-FEB-2002 (first entry)
DE Propionibacterium acnes immunogenic protein #25355.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
DR
XX N-PSDB; AAS59644.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 25654; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in

```

CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 CC Sequence 71 AA;
 CC
 CC Query Match 25.0%; Score 5; DB 4; Length 71;
 CC Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 PRGAP 5
 CC Db 62 PRGAP 66
 CC
 CC RESULT 69
 CC ID ABM60978 standard; protein; 71 AA.
 CC XX
 CC AC ABM60978;
 CC
 CC DT 20-OCT-2003 (first entry)
 CC
 CC DE Propionibacterium acnes predicted ORF-encoded polypeptide #25654.
 CC XX
 CC KW Acne vulgaris; antisporrheic; dermatological; antibacterial;
 CC KW immunostimulant; immune response; vaccine.
 CC XX
 CC OS Propionibacterium acnes.
 CC XX
 CC PN WO2003033515-A1.
 CC XX
 CC PD 24-APR-2003.
 CC
 CC PF 11-OCT-2002; 2002WO-US032727.
 CC
 CC PR 15-OCT-2001; 2001US-00978825.
 CC
 CC PA (CORI-) CORIXA CORP.
 CC
 CC PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JU;
 CC PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 CC PI Barth B, Vallie-Douglass J;
 CC
 CC DR WPI; 2003-381789/36.
 CC DR N-PSDB; ACF64573.
 CC XX
 CC PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
 CC polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 CC or for stimulating an immune response specific for a P. acnes protein.
 CC
 CC PS Example 1; SEQ ID NO 25654; 1481bp; English.
 CC
 CC CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a

CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 CC XX Sequence 71 AA;
 CC
 CC Query Match 25.0%; Score 5; DB 6; Length 71;
 CC Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 PRGAP 5
 CC Db 62 PRGAP 66
 CC
 CC RESULT 70
 CC ID AAM14327
 CC XX AAM14327 standard; protein; 72 AA.
 CC AC AAM14327;
 CC
 CC DT 12-OCT-2001 (first entry)
 CC
 CC DE Peptide #761 encoded by probe for measuring cervical gene expression.
 CC XX
 CC KW Probe: human; microarray; gene expression; cervical epithelial cell;
 CC KW cervical cancer.
 CC XX
 CC OS Homo sapiens.
 CC XX
 CC PN WO200157278-A2.
 CC
 CC PD 09-AUG-2001.
 CC
 CC PF 30-JAN-2001; 2001WO-US000670.
 CC
 CC PR 04-FEB-2000; 2000US-0180312P.
 CC PR 26-MAY-2000; 2000US-0207456P.
 CC PR 30-JUN-2000; 2000US-00608408.
 CC PR 03-AUG-2000; 2000US-00632366.
 CC PR 21-SEP-2000; 2000US-0234687P.
 CC PR 27-SEP-2000; 2000US-0236359P.
 CC PR 04-OCT-2000; 2000GB-00024263.
 CC
 CC PA (MOLE-) MOLECULAR DYNAMICS INC.
 CC XX
 CC PI Penn SG, Hanzel DK, Chen W, Rank DR;
 CC
 CC DR WPI; 2001-488901/53.
 CC
 CC PT Human genome-derived single exon nucleic acid probes useful for analyzing
 CC gene expression in human cervical epithelial cells.
 CC
 CC PS Claim 27; SEQ ID NO 19153; 487bp; English.

XX CC The present invention relates to human single exon nucleic acid probes
 CC (SENPs: see AAI10068-AA128459). The present sequence is a peptide encoded
 CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX CC
 XX SQ Sequence 72 AA;

Query Match 25.0%; Score 5; DB 4; Length 72;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRGAP 5
 Db 16 PRGAP 20
 |||||

RESULT 71
 ABB33274
 ID ABB33274 standard; peptide; 72 AA.
 AC ABB33274;
 XX
 XX 04-FEB-2002 (first entry)
 DT
 XX Peptide #780 encoded by human foetal liver single exon probe.
 DE
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX
 XX Homo sapiens.
 OS
 XX WO200157277-A2.
 FN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US000669.
 PF
 XX 04-FEB-2000; 2000US-0180312P.
 PR
 XX 26-MAY-2000; 2000US-0207456P.
 PR
 XX 30-JUN-2000; 2000US-00608408.
 PR
 XX 03-AUG-2000; 2000US-00632366.
 PR
 XX 21-SEP-2000; 2000US-0234687P.
 PR
 XX 27-SEP-2000; 2000US-0236359P.
 PR
 XX 04-OCT-2000; 2000GB-00024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WPI; 2001-483447/52.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human fetal liver.
 PT
 XX Claim 27; SEQ ID NO 25909; 639pp + Sequence Listing; English.
 PS
 XX The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX CC
 XX SQ Sequence 72 AA;

Query Match 25.0%; Score 5; DB 4; Length 72;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRGAP 5
 Db 16 PRGAP 20
 |||||

RESULT 72
 AAM26737
 ID AAM26737 standard; protein; 72 AA.
 XX
 XX AAM26737;
 AC
 XX 17-OCT-2001 (first entry)
 DT
 XX Peptide #774 encoded by probe for measuring placental gene expression.
 DE
 XX Probe; microarray; human; placenta; antenatal diagnosis;
 XX genetic disorder.
 XX
 XX Homo sapiens.
 OS
 XX WO200157272-A2.
 FN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US000663.
 PF
 XX 04-FEB-2000; 2000US-0180312P.
 PR
 XX 26-MAY-2000; 2000US-0207456P.
 PR
 XX 30-JUN-2000; 2000US-00608408.
 PR
 XX 03-AUG-2000; 2000US-00632366.
 PR
 XX 21-SEP-2000; 2000US-0234687P.
 PR
 XX 27-SEP-2000; 2000US-0236359P.
 PR
 XX 04-OCT-2000; 2000GB-00024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WPI; 2001-488897/53.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human placenta.
 PT
 XX Claim 27; SEQ ID NO 27006; 654pp; English.
 PS
 XX The present invention relates to single exon nucleic acid probes (SENPs:
 CC see AAI31315-AA157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders
 XX
 XX SQ Sequence 72 AA;

Query Match 25.0%; Score 5; DB 4; Length 72;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRGAP 5
 Db 16 PRGAP 20
 |||||

RESULT 73
 ABB28100
 ID ABB28100 standard; peptide; 72 AA.
 XX
 XX ABB28100;
 AC
 XX

DT 01-FEB-2002 (first entry)
 XX Human peptide #751 encoded by breast cell single exon nucleic acid probe.
 DE Human; microarray; single exon probe; gene expression; breast; disease;
 KW cancer.
 KW Homo sapiens.
 XX WO200157271-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US000662.
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-496933/54.
 XX New spatially-addressable set of single exon nucleic acid probes, useful
 PT for measuring gene expression in sample derived from human breast,
 PT comprises number of single exon nucleic acid probes.
 XX Claim 27; SEQ ID NO 11068; 327pp + Sequence Listing; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting the
 CC probes with a collection of detectably labelled nucleic acids derived
 CC from mRNA of human breast, and then measuring the label bound to each
 CC probe of the microarray. The probes are useful for verifying the
 CC expression of regions of genomic DNA predicted to encode proteins. They
 CC are useful for gene discovery, and for determining predisposition and/or
 CC prognosing breast disease. Gene expression analysis is useful for
 CC assessing the toxicity of chemical agents on cells. The microarray of
 CC this invention presents a far greater diversity of probes for measuring
 CC gene expression, with far less bias than expressed sequence tag
 CC microarrays. The method is suitable for rapid production of functional
 CC information from genomic sequence. The present sequence is a peptide
 CC encoded by a single exon nucleic acid probe of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 72 AA;
 SQ
 Query Match 25.0%; Score 5; DB 4; Length 72;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PRGAP 5
 Db 16 PRGAP 20
 RESULT 74
 ABB18736
 ID ABB18736 standard; protein; 72 AA.
 XX ABB18736;
 AC ABB18736;
 XX 23-JAN-2002 (first entry)
 DT
 XX

DE Protein #735 encoded by probe for measuring heart cell gene expression.
 XX Human; gene expression; heart; microarray; vascular system;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease.
 XX Homo sapiens.
 OS WO200157274-A2.
 PN 09-AUG-2001.
 PD 30-JAN-2001; 2001WO-US000666.
 PF 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488899/53.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts.
 PT Claim 15; SEQ ID NO 20506; 530pp; English.
 PS The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting, the
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 72 AA;
 SQ
 Query Match 25.0%; Score 5; DB 4; Length 72;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PRGAP 5
 Db 16 PRGAP 20
 RESULT 75
 AAM66456
 ID AAM66456 standard; protein; 72 AA.
 XX AAM66456;
 AC AAM66456;
 XX 06-NOV-2001 (first entry)
 DT Human bone marrow expressed probe encoded protein SEQ ID NO: 26762.
 DE Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX Homo sapiens.
 OS WO200157276-A2.
 PN

Tue Oct 26 08:02:19 2004

```
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US0000668.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
PR
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PT
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 26762; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
XX Sequence 72 AA;
SQ
Query Match 25.0%; Score 5; DB 4; Length 72;
Best Local Similarity 100.0%; Pred.No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PRGAP 5
Db 16 PRGAP 20
Search completed: October 26, 2004, 07:16:20
Job time : 92 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 07:22:37 ; Search time 65 Seconds
(without alignments)
99.618 Million cell updates/sec

Title: US-10-066-965A-3

Perfect score: 20

Sequence: 1 PRGAPMWRWVQMLETML 20

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1364641 seqs, 323758627 residues

Word size : 0

Total number of hits satisfying Chosen parameters: 1364641

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	14	US-10-066-965A-3
2	20	100.0	20	14	US-10-066-965A-22
3	7	35.0	20	14	US-10-066-965A-4
4	7	35.0	20	14	US-10-066-965A-23
5	6	30.0	34	15	US-10-066-965A-4341
6	6	30.0	253	16	US-10-437-963-114356
7	6	30.0	262	15	US-10-425-114-67611
8	6	30.0	324	16	US-10-479-334-26
9	6	30.0	731	16	US-10-408-765A-2869
10	5	25.0	9	9	US-09-834-765-528
11	5	25.0	9	9	US-09-834-765-634
12	5	25.0	10	9	US-09-834-765-583
13	5	25.0	20	14	US-10-280-066-66
14	5	25.0	20	14	US-10-312-495-25
15	5	25.0	20	14	US-10-312-495-25
16	5	25.0	20	14	US-10-312-495-25
17	5	25.0	20	14	US-10-312-495-25
18	5	25.0	20	14	US-10-312-495-25
19	5	25.0	20	14	US-10-312-495-25
20	5	25.0	20	14	US-10-312-495-25
21	5	25.0	20	14	US-10-312-495-25
22	5	25.0	20	14	US-10-312-495-25
23	5	25.0	20	14	US-10-312-495-25
24	5	25.0	20	14	US-10-312-495-25
25	5	25.0	20	14	US-10-312-495-25
26	5	25.0	20	14	US-10-312-495-25
27	5	25.0	20	14	US-10-312-495-25
28	5	25.0	20	14	US-10-312-495-25
29	5	25.0	20	14	US-10-312-495-25
30	5	25.0	20	14	US-10-312-495-25
31	5	25.0	20	14	US-10-312-495-25
32	5	25.0	20	14	US-10-312-495-25
33	5	25.0	20	14	US-10-312-495-25
34	5	25.0	20	14	US-10-312-495-25
35	5	25.0	20	14	US-10-312-495-25
36	5	25.0	20	14	US-10-312-495-25
37	5	25.0	20	14	US-10-312-495-25
38	5	25.0	20	14	US-10-312-495-25
39	5	25.0	20	14	US-10-312-495-25
40	5	25.0	20	14	US-10-312-495-25
41	5	25.0	20	14	US-10-312-495-25
42	5	25.0	20	14	US-10-312-495-25
43	5	25.0	20	14	US-10-312-495-25
44	5	25.0	20	14	US-10-312-495-25
45	5	25.0	20	14	US-10-312-495-25
46	5	25.0	20	14	US-10-312-495-25
47	5	25.0	20	14	US-10-312-495-25
48	5	25.0	20	14	US-10-312-495-25
49	5	25.0	20	14	US-10-312-495-25
50	5	25.0	20	14	US-10-312-495-25
51	5	25.0	20	14	US-10-312-495-25
52	5	25.0	20	14	US-10-312-495-25
53	5	25.0	20	14	US-10-312-495-25
54	5	25.0	20	14	US-10-312-495-25
55	5	25.0	20	14	US-10-312-495-25
56	5	25.0	20	14	US-10-312-495-25
57	5	25.0	20	14	US-10-312-495-25
58	5	25.0	20	14	US-10-312-495-25
59	5	25.0	20	14	US-10-312-495-25
60	5	25.0	20	14	US-10-312-495-25
61	5	25.0	20	14	US-10-312-495-25
62	5	25.0	20	14	US-10-312-495-25
63	5	25.0	20	14	US-10-312-495-25
64	5	25.0	20	14	US-10-312-495-25
65	5	25.0	20	14	US-10-312-495-25
66	5	25.0	20	14	US-10-312-495-25
67	5	25.0	20	14	US-10-312-495-25
68	5	25.0	20	14	US-10-312-495-25
69	5	25.0	20	14	US-10-312-495-25
70	5	25.0	20	14	US-10-312-495-25
71	5	25.0	20	14	US-10-312-495-25
72	5	25.0	20	14	US-10-312-495-25
73	5	25.0	20	14	US-10-312-495-25
74	5	25.0	20	14	US-10-312-495-25
75	5	25.0	20	14	US-10-312-495-25
76	5	25.0	20	14	US-10-312-495-25
77	5	25.0	20	14	US-10-312-495-25
78	5	25.0	20	14	US-10-312-495-25
79	5	25.0	20	14	US-10-312-495-25
80	5	25.0	20	14	US-10-312-495-25
81	5	25.0	20	14	US-10-312-495-25
82	5	25.0	20	14	US-10-312-495-25
83	5	25.0	20	14	US-10-312-495-25
84	5	25.0	20	14	US-10-312-495-25
85	5	25.0	20	14	US-10-312-495-25
86	5	25.0	20	14	US-10-312-495-25

Sequence 416, App
Sequence 120828, A
Sequence 35124, A
Sequence 45993, A
Sequence 6, Appli
Sequence 38656, A
Sequence 204604, A
Sequence 267231, A
Sequence 13, Appl
Sequence 262093, A
Sequence 1954, Ap
Sequence 219441, A
Sequence 264712, A
Sequence 32143, A
Sequence 152702, A
Sequence 225320, A
Sequence 1576, Ap
Sequence 150750, A
Sequence 34034, A
Sequence 175, App
Sequence 102863, A
Sequence 135152, A
Sequence 144203, A
Sequence 149863, A
Sequence 464, App
Sequence 161301, A
Sequence 105389, A
Sequence 240184, A
Sequence 7816, Ap
Sequence 772, App
Sequence 18176, A
Sequence 47995, A
Sequence 180086, A
Sequence 281783, A
Sequence 102687, A
Sequence 50033, A
Sequence 127248, A
Sequence 115372, A
Sequence 124347, A
Sequence 217028, A
Sequence 7168, Ap
Sequence 135530, A
Sequence 47642, A
Sequence 15, Appl
Sequence 142846, A
Sequence 157410, A
Sequence 159333, A
Sequence 134702, A
Sequence 8165, Ap
Sequence 155904, A
Sequence 303, App
Sequence 104856, A
Sequence 267655, A
Sequence 1456, Ap
Sequence 185681, A
Sequence 187877, A
Sequence 195845, A
Sequence 60687, A
Sequence 46232, A
Sequence 216376, A
Sequence 126718, A
Sequence 280, App
Sequence 144416, A
Sequence 131328, A
Sequence 184086, A
Sequence 188198, A
Sequence 21, Appl
Sequence 154610, A
Sequence 12046, A
Sequence 148533, A
Sequence 149240, A
Sequence 195447, A
Sequence 25, Appl

```

Sequence 212562,
Sequence 131017,
Sequence 57670, A
Sequence 122028,
Sequence 259958,
Sequence 1953, Ap
Sequence 143716,
Sequence 161234,
Sequence 157405,
Sequence 182171,
Sequence 118153,
Sequence 153955,
Sequence 281122,
Sequence 128134,

87 5 25.0 114 15 US-10-424-599-212562
88 5 25.0 114 16 US-10-437-963-131017
89 5 25.0 114 16 US-10-767-701-57670
90 5 25.0 116 16 US-10-437-963-122028
91 5 25.0 117 15 US-10-424-599-259958
92 5 25.0 119 15 US-10-264-237-1953
93 5 25.0 120 16 US-10-437-963-143716
94 5 25.0 121 16 US-10-437-963-161234
95 5 25.0 121 16 US-10-437-963-157405
96 5 25.0 122 16 US-10-437-963-182171
97 5 25.0 123 16 US-10-437-963-118153
98 5 25.0 123 16 US-10-437-963-153955
99 5 25.0 126 15 US-10-424-599-281122
100 5 25.0 126 16 US-10-437-963-128134

ALIGNMENTS

RESULT 1
US-10-066-965A-3
; Sequence 3, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-3

Query Match 100.0%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.1e-15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRGAPMMRWVCQMLETMFL 20
Db 1 PRGAPMMRWVCQMLETMFL 20

RESULT 2
US-10-066-965A-22
; Sequence 22, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-22

Query Match 100.0%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.1e-15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRGAPMMRWVCQMLETMFL 20
Db 1 PRGAPMMRWVCQMLETMFL 20

RESULT 3
US-10-066-965A-4
; Sequence 4, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-4

Query Match 35.0%; Score 7; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRGAPMW 7
Db 1 PRGAPMW 7

RESULT 4
US-10-066-965A-23
; Sequence 23, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-23

Query Match 35.0%; Score 7; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRGAPMW 7
Db 1 PRGAPMW 7
```

```
RESULT 5
US-10-264-049-4341
; Sequence 4341, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P413P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4341
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-4341

Query Match          30.0%; Score 6; DB 15; Length 34;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMFL 20
    |||||
Db 2 LETMFL 7

RESULT 6
US-10-437-963-114356
; Sequence 114356, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 114356
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_18055C.1.pep
US-10-437-963-114356

Query Match          30.0%; Score 6; DB 16; Length 253;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GAPMM 8
    |||||
Db 18 GAPMM 23

RESULT 7
US-10-425-114-67611
; Sequence 67611, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P413P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4341
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-4341

Query Match          30.0%; Score 6; DB 15; Length 34;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMFL 20
    |||||
Db 2 LETMFL 7

RESULT 8
US-10-479-334-26
; Sequence 26, Application US/10479334
; Publication No. US20040137460A1
; GENERAL INFORMATION:
; APPLICANT: YAMANAKA, Shinya
; APPLICANT: KAIHO, Hideko
; TITLE OF INVENTION: GENES WITH ES CELL-SPECIFIC EXPRESSION
; FILE REFERENCE: 225475
; CURRENT APPLICATION NUMBER: US/10/479,334
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: PCT/JPO2/05350
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: JP 2001-165927
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-479-334-26

Query Match          30.0%; Score 6; DB 16; Length 324;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMFL 20
    |||||
Db 210 LETMFL 215

RESULT 9
US-10-408-765A-2869
; Sequence 2869, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
```

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 ; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
 ; FILE REFERENCE: 660088.465
 ; CURRENT APPLICATION NUMBER: US/10/408,765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2869
 ; LENGTH: 731
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-408-765A-2869

Query Match 30.0%; Score 6; DB 16; Length 731;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAPM 6
 Db 542 PRGAPM 547

RESULT 10
 US-09-834-765-528
 ; Sequence 528, Application US/09834765
 ; Patent No. US20020055478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mary Faris
 ; APPLICANT: Pia M. Challita-Eid
 ; APPLICANT: Arthur B. Raitano
 ; APPLICANT: Steve Chappell Mitchell
 ; APPLICANT: Daniel E.H. Afar
 ; APPLICANT: Aya Jakobovits
 ; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
 ; TITLE OF INVENTION: AND DETECTION OF CANCER
 ; FILE REFERENCE: 129.6USU1
 ; CURRENT APPLICATION NUMBER: US/09/834,765
 ; CURRENT FILING DATE: 2001-09-21
 ; PRIOR APPLICATION NUMBER: 60/197,647
 ; PRIOR FILING DATE: 2000-04-12
 ; NUMBER OF SEQ ID NOS: 770
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 528
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-834-765-528

Query Match 25.0%; Score 5; DB 9; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.2e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 Db 2 PRGAP 6

RESULT 11
 US-09-834-765-634
 ; Sequence 634, Application US/09834765
 ; Patent No. US20020055478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mary Faris
 ; APPLICANT: Pia M. Challita-Eid
 ; APPLICANT: Arthur B. Raitano
 ; APPLICANT: Steve Chappell Mitchell
 ; APPLICANT: Daniel E.H. Afar
 ; APPLICANT: Aya Jakobovits
 ; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
 ; TITLE OF INVENTION: AND DETECTION OF CANCER
 ; FILE REFERENCE: 129.6USU1
 ; CURRENT APPLICATION NUMBER: US/09/834,765
 ; CURRENT FILING DATE: 2001-09-21

; PRIOR APPLICATION NUMBER: 60/197,647
 ; PRIOR FILING DATE: 2000-04-12
 ; NUMBER OF SEQ ID NOS: 770
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 634
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-834-765-634

Query Match 25.0%; Score 5; DB 9; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.2e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 Db 2 PRGAP 6

RESULT 12
 US-09-834-765-583
 ; Sequence 583, Application US/09834765
 ; Patent No. US20020055478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mary Faris
 ; APPLICANT: Pia M. Challita-Eid
 ; APPLICANT: Arthur B. Raitano
 ; APPLICANT: Steve Chappell Mitchell
 ; APPLICANT: Daniel E.H. Afar
 ; APPLICANT: Aya Jakobovits
 ; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
 ; TITLE OF INVENTION: AND DETECTION OF CANCER
 ; FILE REFERENCE: 129.6USU1
 ; CURRENT APPLICATION NUMBER: US/09/834,765
 ; CURRENT FILING DATE: 2001-09-21
 ; PRIOR APPLICATION NUMBER: 60/197,647
 ; PRIOR FILING DATE: 2000-04-12
 ; NUMBER OF SEQ ID NOS: 770
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 583
 ; LENGTH: 10
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-834-765-583

Query Match 25.0%; Score 5; DB 9; Length 10;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 Db 2 PRGAP 6

RESULT 13
 US-10-280-066-66
 ; Sequence 66, Application US/10280066
 ; Publication No. US20030180718A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pallatla, Renuka C.
 ; APPLICANT: Brissette, Renee
 ; APPLICANT: Spruyt, Michael
 ; APPLICANT: Dedova, Olga
 ; APPLICANT: Blume, Arthur J.
 ; APPLICANT: Prendergast, John
 ; APPLICANT: Goldstein, Neil I.
 ; TITLE OF INVENTION: TARGET SPECIFIC SCREENING AND ITS USE FOR IDENTIFYING TARGET BIND
 ; FILE REFERENCE: 2598-4009US1
 ; CURRENT APPLICATION NUMBER: US/10/280,066
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 60/345,471
 ; PRIOR FILING DATE: 2001-10-24
 ; NUMBER OF SEQ ID NOS: 537

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 66

; LENGTH: 20

; TYPE: PRT

; ORGANISM: Escherichia coli

; FEATURE:

; NAME/KEY: MISC FEATURE

; OTHER INFORMATION: DGI-2-20R-4-G114

US-10-280-066-66

Query Match 25.0%; Score 5; DB 14; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 MMRW 10

|||||

DB 8 MMRW 12

RESULT 14

US-10-231-417-416

; Sequence 416, Application US/10231417

; Publication No. US20030176681A1

; GENERAL INFORMATION:

; APPLICANT: Feng et al.

; TITLE OF INVENTION: 148 Human Secreted Proteins

; FILE REFERENCE: P2019P1

; CURRENT APPLICATION NUMBER: US/10/231,417

; CURRENT FILING DATE: 2002-08-30

; PRIOR APPLICATION NUMBER: US/09/296,622

; PRIOR FILING DATE: 1999-04-23

; NUMBER OF SEQ ID NOS: 619

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 416

; LENGTH: 28

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-231-417-416

Query Match 25.0%; Score 5; DB 14; Length 28;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CQMLE 16

|||||

DB 3 CQMLE 7

RESULT 15

US-10-437-963-120828

; Sequence 120828, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 120828

; LENGTH: 44

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_23912C.1.pep

US-10-437-963-120828

Query Match

Best Local Similarity 100.0%; Score 5; DB 16; Length 44;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCQML 15

|||||

DB 10 VCQML 14

RESULT 16

US-09-864-761-35124

; Sequence 35124, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aecmica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 35124

; LENGTH: 46

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC007688.15

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3

; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8

; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.1

```
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.99
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EST_HUMAN HIT: R77154.1, EVALUE 1.00e-08
US-09-864-761-35124
```

```
Query Match      25.0%; Score 5; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 RGAPM 6
        |||||
Db      20 RGAPM 24
```

```
RESULT 17
US-09-864-761-45993
; Sequence 45993, Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45993
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; OTHER INFORMATION: MAP TO AC005323.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2e+02
; OTHER INFORMATION: EST_HUMAN HIT: BF339300.1, EVALUE 4.00e+00
US-09-864-761-45993
```

```
Query Match      25.0%; Score 5; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      13 QMLET 17
        |||||
Db      32 QMLET 36
```

```
RESULT 18
US-10-815-514-6
; Sequence 6, Application US/10815514
; Publication No. US20040204361A1
; GENERAL INFORMATION:
; APPLICANT: Rothman, James
; APPLICANT: Mayhew, Mark
; APPLICANT: Hoe, Mee
; TITLE OF INVENTION: KDEL RECEPTOR INHIBITORS
; FILE REFERENCE: 31488
; CURRENT APPLICATION NUMBER: US/10/815,514
; CURRENT FILING DATE: 2004-03-31
; PRIOR APPLICATION NUMBER: US/09/696,872
; PRIOR FILING DATE: 2000-10-26
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-815-514-6
```

```
Query Match      25.0%; Score 5; DB 17; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      16 ETMFL 20
        |||||
Db      31 ETMFL 35
```

```
RESULT 19
US-09-864-761-38656
; Sequence 38656, Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
```

;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 38656
;; LENGTH: 47
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006024.1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
;; OTHER INFORMATION: EST_HUMAN HIT: BF241410.1, EVALUATE 6.00e-01
US-09-864-761-38656

Query Match 25.0%; Score 5; DB 9; Length 47;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
DB 13 PRGAP 17

RESULT 20
US-10-437-963-204604
; Sequence 204604, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14

;; NUMBER OF SEQ ID NOS: 204966
;; SEQ ID NO 204604
;; LENGTH: 50
;; TYPE: PRT
;; ORGANISM: Oryza sativa
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT4530_99675C.1.pep
US-10-437-963-204604

Query Match 25.0%; Score 5; DB 16; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
DB 44 PRGAP 48

RESULT 21
US-10-424-599-267231
; Sequence 267231, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 267231
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_8332C.1.pep
US-10-424-599-267231

Query Match 25.0%; Score 5; DB 15; Length 51;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCQML 15
DB 41 VCQML 45

RESULT 22
US-10-152-031-13
; Sequence 13, Application US/10152031
; Publication No. US20030044825A1
; GENERAL INFORMATION:
; APPLICANT: Imai, Yuji
; APPLICANT: AKATSUKA, Hiroyuki
; APPLICANT: KAWAI, Eri
; APPLICANT: OMORI, Kenji
; APPLICANT: YANAKA, No. US20030044825A1iyuki
; APPLICANT: SAKURAI, Naoki
; TITLE OF INVENTION: Bone Metabolism Related Protein and Gene Thereof
; FILE REFERENCE: 0283-0163P
; CURRENT APPLICATION NUMBER: US/10/152,031
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,318
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: JP 318226/1998
; PRIOR FILING DATE: 2000-05-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 54

; TYPE: PRT
; ORGANISM: Mus musculus
US-10-152-031-13

Query Match 25.0%; Score 5; DB 14; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
| | | | |
DB 4 RGAPM 8

RESULT 23

US-10-424-599-262093
; Sequence 262093, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 262093
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_78692C.1.pep
US-10-424-599-262093

Query Match 25.0%; Score 5; DB 15; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
| | | | |
DB 13 PRGAP 17

RESULT 24

US-09-867-550-1954
; Sequence 1954, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1954
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-1954

Query Match 25.0%; Score 5; DB 9; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCQML 15
| | | | |
DB 22 VCQML 26

RESULT 25

US-10-424-599-219441
; Sequence 219441, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 219441
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_40181C.1.pep
US-10-424-599-219441

Query Match 25.0%; Score 5; DB 15; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETM 18
| | | | |
DB 9 MLETM 13

RESULT 26

US-10-424-599-264712
; Sequence 264712, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 264712
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_81055C.1.pep
US-10-424-599-264712

Query Match 25.0%; Score 5; DB 15; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
| | | | |
DB 17 LETMF 21

RESULT 27

US-10-029-386-32143
; Sequence 32143, Application US/10029386


```

; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32143
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO Z98883.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: SWISSPROT HIT: O43451, EVALUE 2.00e-03
US-10-029-386-32143

Query Match 25.0%; Score 5; DB 14; Length 63;
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 1 PRGAP 5
DB 15 PRGAP 19

RESULT 28
US-10-424-599-152702
; Sequence 152702, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 152702
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108914C.1.pep
US-10-424-599-152702

Query Match 25.0%; Score 5; DB 15; Length 65;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 6 MMRW 10
DB 28 MMRW 32

RESULT 29
US-10-424-599-225320
; Sequence 225320, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 152702
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_45494C.1.pep
US-10-424-599-225320

Query Match 25.0%; Score 5; DB 15; Length 66;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 15 LETMF 19
DB 3 LETMF 7

RESULT 30
US-09-867-550-1576
; Sequence 1576, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells an
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1576
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Clone ID: PAT_MRT3847_45494C.1.pep
US-09-867-550-1576

Query Match 25.0%; Score 5; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 1 PRGAP 5
DB 40 PRGAP 44

RESULT 31
US-10-424-599-150750
; Sequence 150750, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 152702
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108914C.1.pep
US-10-424-599-152702

Query Match 25.0%; Score 5; DB 15; Length 65;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 6 MMRW 10
DB 28 MMRW 32

RESULT 29
US-10-424-599-225320
; Sequence 225320, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 152702
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_45494C.1.pep
US-10-424-599-225320

Query Match 25.0%; Score 5; DB 15; Length 66;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 15 LETMF 19
DB 3 LETMF 7

RESULT 30
US-09-867-550-1576
; Sequence 1576, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells an
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1576
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Clone ID: PAT_MRT3847_45494C.1.pep
US-09-867-550-1576

Query Match 25.0%; Score 5; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 1 PRGAP 5
DB 40 PRGAP 44

RESULT 31
US-10-424-599-150750
; Sequence 150750, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 152702
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108914C.1.pep
US-10-424-599-152702

Query Match 25.0%; Score 5; DB 15; Length 65;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 6 MMRW 10
DB 28 MMRW 32

RESULT 29
US-10-424-599-225320
; Sequence 225320, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 152702
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_45494C.1.pep
US-10-424-599-225320

Query Match 25.0%; Score 5; DB 15; Length 66;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 15 LETMF 19
DB 3 LETMF 7

RESULT 30
US-09-867-550-1576
; Sequence 1576, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells an
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1576
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Clone ID: PAT_MRT3847_45494C.1.pep
US-09-867-550-1576

Query Match 25.0%; Score 5; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;
```

; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 150750
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(68)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_107151C.1.pep
US-10-424-599-150750

Query Match 25.0%; Score 5; DB 15; Length 68;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 MLETM 18
Db 21 MLETM 25

RESULT 32

US-09-864-761-34034
; Sequence 34034, Application US/09864761
; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensteng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aecmica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34034
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AB023057.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.87
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
; OTHER INFORMATION: EST HUMAN HIT: BE314635.1, EVALUUE 5.00e-09
; OTHER INFORMATION: EST_HUMAN HIT: AU139120.1, EVALUUE 1.00e-15
US-09-864-761-34034

Query Match 25.0%; Score 5; DB 9; Length 72;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRGAP 5

Db 16 PRGAP 20

RESULT 33

US-10-001-879-175

; Sequence 175, Application US/10001879

; Publication No. US20020127237A1

; GENERAL INFORMATION:

; APPLICANT: Salceda, Susana

; APPLICANT: Macina, Roberto

; APPLICANT: Recipon, Herve

; APPLICANT: Cafferkey, Robert

; APPLICANT: Ali, Shujath

; APPLICANT: Sun, Yongming

; APPLICANT: Liu, Chenghua

; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and

; FILE REFERENCE: DEX-0281

; CURRENT APPLICATION NUMBER: US/10/001,879

; CURRENT FILING DATE: 2001-11-20

; PRIOR APPLICATION NUMBER: 60/252,188

; PRIOR FILING DATE: 2000-11-21

; NUMBER OF SEQ ID NOS: 201

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 175

; LENGTH: 72

; TYPE: PRT

; ORGANISM: Homo sapien

US-10-001-879-175

Query Match 25.0%; Score 5; DB 13; Length 72;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RGAPM 6

Db 10 RGAPM 14

RESULT 34

US-10-437-963-102863

; Sequence 102863, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 102863
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_100348C.1.pep
US-10-437-963-102863

Query Match      25.0%; Score 5; DB 16; Length 74;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
Db      30 PRGAP 34

RESULT 35
US-10-437-963-135152
; Sequence 135152, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 135152
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_36855C.1.pep
US-10-437-963-135152

Query Match      25.0%; Score 5; DB 16; Length 78;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
Db      23 PRGAP 27

RESULT 36
US-10-424-599-144203
; Sequence 144203, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
```

```
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 144203
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101228C.1.pep
US-10-424-599-144203

Query Match      25.0%; Score 5; DB 15; Length 79;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
Db      43 PRGAP 47

RESULT 37
US-10-424-599-149863
; Sequence 149863, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 149863
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106348C.1.pep
US-10-424-599-149863

Query Match      25.0%; Score 5; DB 15; Length 79;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      13 QMLET 17
Db      20 QMLET 24

RESULT 38
US-09-922-261-464
; Sequence 464, Application US/09922261
; Patent No. US20020111471A1
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; FILE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
```

```

; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 105389
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_102638C.1.pep
US-10-437-963-105389

Query Match          25.0%; Score 5; DB 16; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
        |||||
DB      24 PRGAP 28

RESULT 41
US-10-424-599-240184
; Sequence 240184, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 240184
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_58913C.1.pep
US-10-424-599-240184

Query Match          25.0%; Score 5; DB 15; Length 83;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
        |||||
DB      55 PRGAP 59

RESULT 42
US-10-106-698-7816
; Sequence 7816, Application US/101066698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 7816
; LENGTH: 84
; TYPE: PRT

```

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (2)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC FEATURE
LOCATION: (14)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC FEATURE
LOCATION: (31)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC FEATURE
LOCATION: (39)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC FEATURE
LOCATION: (41)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC FEATURE
LOCATION: (43)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC FEATURE
LOCATION: (65)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC FEATURE
LOCATION: (72)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC FEATURE
LOCATION: (77)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC FEATURE
LOCATION: (82)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-7816

Query Match 25.0%; Score 5; DB 14; Length 84;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5

Db 6 PRGAP 10

RESULT 43
US-09-867-550-772
Sequence 772, Application US/09867550
Patent No. US20020082206A1
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Fuad,
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
APPLICANT: Topper, James
TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
FILE REFERENCE: 21402-013 (Cura-313)
CURRENT APPLICATION NUMBER: US/09/867,550
CURRENT FILING DATE: 2001-09-20
PRIOR APPLICATION NUMBER: USSN 60/208,427
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 2125
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 772
LENGTH: 85
TYPE: PRT
ORGANISM: Homo sapiens
US-09-867-550-772

Query Match 25.0%; Score 5; DB 9; Length 85;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5

Db 38 PRGAP 42

RESULT 44
US-10-437-963-183176
Sequence 183176, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 183176
LENGTH: 85
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_80293C.1.pap
US-10-437-963-183176

Query Match 25.0%; Score 5; DB 16; Length 85;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5

Db 57 PRGAP 61

RESULT 45
US-09-864-761-47995
Sequence 47995, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 47995
LENGTH: 86
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AB023057.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.66
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.68
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.94
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.67
OTHER INFORMATION: EST_HUMAN HIT: BE888511.1, EVALUATE 1.00e-10
US-09-864-761-47995

Query Match 25.0%; Score 5; DB 9; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 30 PRGAP 34

RESULT 46
US-10-424-599-180086
; Sequence 180086, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 180086
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_133632C.1.pap
US-10-424-599-180086

Query Match 25.0%; Score 5; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PMWNR 9
|||||
Db 17 PMWNR 21

RESULT 47
US-10-424-599-281783
; Sequence 281783, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 281783
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_96470C.1.pap
US-10-424-599-281783

Query Match 25.0%; Score 5; DB 15; Length 87;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 61 PRGAP 65

RESULT 48
US-10-437-963-102687
; Sequence 102687, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 102687
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_100186C.1.pap
US-10-437-963-102687

Query Match 25.0%; Score 5; DB 16; Length 88;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 71 PRGAP 75

RESULT 49
US-10-767-701-50033
; Sequence 50033, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 50033
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3478-035-P1-K1-H10.ppep
US-10-767-701-50033

Query Match 25.0%; Score 5; DB 16; Length 88;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCQML 15
Db 53 VCQML 57

RESULT 50
US-10-437-963-127248
; Sequence 127248, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 127248
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29718C.1.ppep
US-10-437-963-127248

Query Match 25.0%; Score 5; DB 16; Length 89;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 81 PRGAP 85

RESULT 51
US-10-437-963-115372
; Sequence 115372, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 115372
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_18976C.1.ppep
US-10-437-963-115372

Query Match 25.0%; Score 5; DB 16; Length 92;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 36 PRGAP 40

RESULT 52
US-10-437-963-124347
; Sequence 124347, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 124347
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_27095C.1.ppep
US-10-437-963-124347

Query Match 25.0%; Score 5; DB 16; Length 92;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 35 PRGAP 39

RESULT 53
US-10-424-599-217028
; Sequence 217028, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B

```
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 217028
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_38003C.1.pep
US-10-424-599-217028

Query Match      25.0%; Score 5; DB 15; Length 93;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      15 LETMP 19
Db      54 LETMP 58

RESULT 54
US-09-864-408A-7168
; Sequence 7168, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 9068
; SEQ ID NO 7168
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-09-864-408A-7168

Query Match      25.0%; Score 5; DB 11; Length 94;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PRGAP 5
Db      69 PRGAP 73

RESULT 55
US-10-437-963-153530
; Sequence 153530, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
```

```
; SEQ ID NO 153530
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_53477C.1.pep
US-10-437-963-153530

Query Match      25.0%; Score 5; DB 16; Length 95;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PRGAP 5
Db      42 PRGAP 46

RESULT 56
US-10-767-701-47642
; Sequence 47642, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 47642
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; NAME/KEY: unsure
; LOCATION: (1)..(95)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3476-034-P1-K1-G12.pep
US-10-767-701-47642

Query Match      25.0%; Score 5; DB 16; Length 95;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      16 ETMFL 20
Db      88 ETMFL 92

RESULT 57
US-09-905-243-15
; Sequence 15, Application US/09905243
; Patent No. US20020062009A1
; GENERAL INFORMATION:
; APPLICANT: Taylor, Alexander H
; TITLE OF INVENTION: Monoclonal Antibodies with Reduced
; TITLE OF INVENTION: Immunogenicity
; FILE REFERENCE: P50770
; CURRENT APPLICATION NUMBER: US/09/905,243
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/300,970
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: DOMAIN
```



```
; LOCATION: (31)...(35)
; OTHER INFORMATION: CDRI
; NAME/KEY: DOMAIN
; LOCATION: (50)...(66)
; OTHER INFORMATION: CDRII
US-09-905-243-15

Query Match      25.0%; Score 5; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WVCQM 14
Db 36 WVCQM 40

RESULT 58
US-10-424-599-142846
; Sequence 142846, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 142846
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100003C.1.pep
US-10-424-599-142846

Query Match      25.0%; Score 5; DB 15; Length 96;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCQML 15
Db 90 VCQML 94

RESULT 59
US-10-437-963-157410
; Sequence 157410, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 157410
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_56985C.1.pep
US-10-437-963-157410

Query Match      25.0%; Score 5; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 60 PRGAP 64

RESULT 60
US-10-424-599-159333
; Sequence 159333, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 159333
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_114899C.1.pep
US-10-424-599-159333

Query Match      25.0%; Score 5; DB 15; Length 97;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 8 PRGAP 12

RESULT 61
US-10-437-963-134702
; Sequence 134702, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 134702
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_3644C.1.pep
US-10-437-963-134702

Query Match      25.0%; Score 5; DB 16; Length 97;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 PRGAP 5
|||||
Db 40 PRGAP 44

RESULT 62

US-10-156-761-8165
; Sequence 8165, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8165
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8165

Query Match 25.0%; Score 5; DB 14; Length 98;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETM 18
|||||
Db 1 MLETM 5

RESULT 63

US-10-424-599-155904
; Sequence 155904, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 155904
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_111801C.1.psp
US-10-424-599-155904

Query Match 25.0%; Score 5; DB 15; Length 98;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCOML 15
|||||
Db 2 VCOML 6

RESULT 64

US-09-986-480-303
; Sequence 303, Application US/09986480
; Publication No. US20030027999A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 143 Human Secreted Proteins
; FILE REFERENCE: PS500P1
; CURRENT APPLICATION NUMBER: US/09/986,480
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/12788
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/134,068
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 303
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-986-480-303

Query Match 25.0%; Score 5; DB 10; Length 100;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 63 PRGAP 67

RESULT 65

US-10-437-963-104856
; Sequence 104856, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 104856
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_10214C.1.psp
US-10-437-963-104856

Query Match 25.0%; Score 5; DB 16; Length 100;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 69 PRGAP 73

RESULT 66

US-10-424-599-267655
; Sequence 267655, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 267655
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_83715C.1.pep
US-10-424-599-267655

Query Match 25.0%; Score 5; DB 15; Length 101;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CQML 16
|||
Db 89 CQML 93

RESULT 67

US-10-264-237-1456
; Sequence 1456, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1456
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (42)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (64)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-1456

Query Match 25.0%; Score 5; DB 15; Length 103;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCQML 15
|||
Db 25 VCQML 29

RESULT 68

US-10-437-963-185681
; Sequence 185681, Application US/10437963

; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 185681
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_82553C.1.pep
US-10-437-963-185681

Query Match 25.0%; Score 5; DB 16; Length 103;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETMFL 20
|||
Db 99 ETMFL 103

RESULT 69

US-10-437-963-187877
; Sequence 187877, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 187877
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84537C.1.pep
US-10-437-963-187877

Query Match 25.0%; Score 5; DB 16; Length 103;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||
Db 24 PRGAP 28

RESULT 70

US-10-437-963-195845
; Sequence 195845, Application US/10437963
; Publication No. US20040123343A1

GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195845
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(103)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91754C.1.pap
US-10-437-963-195845

Query Match 25.0%; Score 5; DB 16; Length 103;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 77 PRGAP 81

RESULT 71
US-10-767-701-60687
; Sequence 60687, Application US/107677701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 60687
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 9281056.pap
US-10-767-701-60687

Query Match 25.0%; Score 5; DB 16; Length 103;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
Db 91 RGAPM 95

RESULT 72
US-10-425-114-46232
; Sequence 46232, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46232
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701130375_FLL.pap
US-10-425-114-46232

Query Match 25.0%; Score 5; DB 15; Length 104;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
Db 79 LETMF 83

RESULT 73
US-10-424-599-216376
; Sequence 216376, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 216376
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(105)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_37416C.1.pap
US-10-424-599-216376

Query Match 25.0%; Score 5; DB 15; Length 105;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
Db 73 RGAPM 77

RESULT 74
US-10-437-963-126718
; Sequence 126718, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126718
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29238C.1.pep
US-10-437-963-126718

Query Match 25.0%; Score 5; DB 16; Length 105;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 46 PRGAP 50

RESULT 75
US-09-833-245-280
; Sequence 280, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 280
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-280

Query Match 25.0%; Score 5; DB 11; Length 107;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 81 PRGAP 85

Search completed: October 26, 2004, 07:47:56
Job time : 66 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 07:06:01 ; Search time 22.25 Seconds
(without alignments)
59.612 Million cell updates/sec

Title: US-10-066-965A-3

Perfect score: 20

Sequence: 1 PRGAPMMRWVCQMLTMEFL 20

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 478139 seqs, 66318000 residues

Word size : 0

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A COMB.pep:*
- 2: /cgn2_6/prodata/1/iaa/5B COMB.pep:*
- 3: /cgn2_6/prodata/1/iaa/6A COMB.pep:*
- 4: /cgn2_6/prodata/1/iaa/6B COMB.pep:*
- 5: /cgn2_6/prodata/1/iaa/PCUS COMB.pep:*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query	ID	Description
1	6	30.0	4	US-08-948-113D-34	Sequence 34, Appl
2	5	25.0	1	US-07-972-032-15	Sequence 15, Appl
3	5	25.0	9	US-08-642-255-15	Sequence 15, Appl
4	5	25.0	46	US-09-124-671-6	Sequence 6, Appl
5	5	25.0	74	US-09-134-001C-3074	Sequence 3074, Ap
6	5	25.0	74	US-09-134-001C-4851	Sequence 4851, Ap
7	5	25.0	77	US-09-270-767-36912	Sequence 36912, A
8	5	25.0	77	US-09-270-767-52129	Sequence 52129, A
9	5	25.0	80	US-09-461-697-464	Sequence 461, App
10	5	25.0	91	US-09-252-991A-24182	Sequence 24182, A
11	5	25.0	91	US-09-270-767-38667	Sequence 38667, A
12	5	25.0	91	US-09-270-767-53884	Sequence 53884, A
13	5	25.0	105	US-09-489-039A-13486	Sequence 13486, A
14	5	25.0	109	US-09-124-671-21	Sequence 21, Appl
15	5	25.0	119	US-09-489-039A-11951	Sequence 11951, A
16	5	25.0	125	US-09-270-767-34903	Sequence 34903, A
17	5	25.0	125	US-09-270-767-50120	Sequence 50120, A
18	5	25.0	133	US-09-252-991A-30594	Sequence 30594, A
19	5	25.0	139	US-09-270-767-39768	Sequence 39768, A
20	5	25.0	139	US-09-270-767-54985	Sequence 54985, A
21	5	25.0	160	US-09-252-991A-23275	Sequence 23275, A
22	5	25.0	163	US-09-134-000C-3710	Sequence 3710, Ap
23	5	25.0	169	US-09-489-039A-8944	Sequence 8944, Ap
24	5	25.0	183	US-09-252-991A-21850	Sequence 21850, A
25	5	25.0	185	US-09-252-991A-32100	Sequence 32100, A
26	5	25.0	199	US-09-270-767-47191	Sequence 47191, A
27	5	25.0	212	US-09-252-991A-26521	Sequence 26521, A

220	4	US-09-270-767-60336	Sequence 60336, A
237	4	US-09-252-991A-23304	Sequence 23304, A
239	3	US-09-134-001C-4165	Sequence 4165, Ap
242	4	US-09-252-991A-28343	Sequence 28343, A
248	1	US-08-426-819A-37	Sequence 37, Appl
257	4	US-09-248-796A-27518	Sequence 27518, A
259	4	US-09-252-991A-21547	Sequence 21547, A
276	4	US-09-710-279-1482	Sequence 1482, Ap
279	1	US-08-312-387B-8	Sequence 8, Appl
279	1	US-08-683-426-8	Sequence 8, Appl
279	1	US-08-683-458-8	Sequence 8, Appl
279	3	US-08-878-360-8	Sequence 8, Appl
279	3	US-09-333-412-8	Sequence 8, Appl
279	4	US-10-007-267-8	Sequence 8, Appl
280	1	US-08-312-387B-6	Sequence 6, Appl
280	1	US-08-683-426-6	Sequence 6, Appl
280	1	US-08-683-458-6	Sequence 6, Appl
280	2	US-08-878-360-6	Sequence 6, Appl
280	3	US-08-478-140B-6	Sequence 6, Appl
280	3	US-09-333-412-6	Sequence 6, Appl
280	3	US-09-338-943-6	Sequence 6, Appl
280	4	US-10-007-267-6	Sequence 6, Appl
283	4	US-09-904-615-106	Sequence 106, App
288	4	US-09-252-991A-27676	Sequence 27676, A
289	4	US-09-252-991A-25496	Sequence 25496, A
297	4	US-09-328-352-7849	Sequence 7849, Ap
304	4	US-09-583-110-5101	Sequence 5101, Ap
308	2	US-08-789-609A-2	Sequence 2, Appl
308	3	US-09-108-517-2	Sequence 16, Appl
308	4	US-10-138-701-16	Sequence 4824, Ap
311	3	US-09-134-001C-4824	Sequence 19219, A
315	4	US-09-252-991A-19219	Sequence 19369, A
327	4	US-09-252-991A-19369	Sequence 32186, A
330	4	US-09-252-991A-32186	Sequence 19891, A
339	4	US-09-252-991A-19891	Sequence 46634, A
354	4	US-09-270-767-46634	Sequence 20554, A
358	4	US-09-252-991A-20554	Sequence 17858, A
363	4	US-09-252-991A-17858	Sequence 31139, A
374	4	US-09-252-991A-31139	Sequence 23047, A
378	4	US-09-252-991A-23047	Sequence 19344, A
394	4	US-09-252-991A-19344	Sequence 26116, A
395	4	US-09-252-991A-26116	Sequence 30428, A
420	4	US-09-328-352-5815	Sequence 5815, Ap
421	4	US-08-396-218-2	Sequence 2, Appl
422	1	US-08-760-116-2	Sequence 2, Appl
422	1	US-09-096-982-5	Sequence 5, Appl
422	2	US-08-653-650A-5	Sequence 5, Appl
425	4	US-09-252-991A-19054	Sequence 19054, A
428	4	US-09-252-991A-27023	Sequence 27023, A
437	4	US-09-252-991A-25619	Sequence 25619, A
438	4	US-09-252-991A-32407	Sequence 32407, A
439	4	US-09-252-991A-21361	Sequence 21361, A
443	2	US-09-096-982-9	Sequence 9, Appl
443	2	US-08-653-650A-9	Sequence 9, Appl
445	4	US-09-252-991A-23505	Sequence 23505, A
450	4	US-09-252-991A-18255	Sequence 18255, A
451	4	US-09-252-991A-27602	Sequence 27602, A
453	4	US-09-328-352-4242	Sequence 4242, Ap
469	4	US-09-252-991A-25438	Sequence 25438, A
474	2	US-09-096-982-8	Sequence 8, Appl
474	2	US-08-653-650A-8	Sequence 8, Appl
474	3	US-08-729-416C-11	Sequence 11, Appl
474	3	US-08-807-342B-2	Sequence 2, Appl
474	4	US-09-433-353-11	Sequence 11, Appl
486	4	US-10-140-002-58	Sequence 58, Appl
499	4	US-09-252-991A-31532	Sequence 31532, A
499	4	US-09-543-681A-6892	Sequence 6892, Ap
516	4	US-09-252-991A-29502	Sequence 29502, A
529	4	US-09-252-991A-24711	Sequence 24711, A
538	4	US-09-616-289-43	Sequence 43, Appl
539	4	US-10-140-002-140	Sequence 140, App
553	4	US-09-252-991A-17429	Sequence 17429, A

ALIGNMENTS

```
RESULT 1
US-08-948-113D-34
; Sequence 34, Application US/08948113D
; Patent No. 6482937
; GENERAL INFORMATION:
; APPLICANT: Baetscher, Manfred W.
; APPLICANT: Akiyoshi, Donna E.
; APPLICANT: Kaplan, Ruth A.
; TITLE OF INVENTION: Pluripotent Porcine Cells, Genetically Modified Porcine
; TITLE OF INVENTION: Cells and Pigs for Use in Said Method, Transgenic Figs
; FILE REFERENCE: 61750-309
; CURRENT APPLICATION NUMBER: US/08/948,113D
; CURRENT FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Mus sp.
US-08-948-113D-34

Query Match          30.0%; Score 6; DB 4; Length 324;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      15 LETMFL 20
DB      210 LETMFL 215

RESULT 2
US-07-972-032-15
; Sequence 15, Application US/07972032
; Patent No. 5496712
; GENERAL INFORMATION:
; APPLICANT: Cappello, Joseph
; APPLICANT: Ferrari, Franco A.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT COLLAGEN-LIKE
; TITLE OF INVENTION: PROTEIN POLYMERS
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bertram I. Rowland
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: CA 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/972,032
; FILING DATE: 19921105
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/791,960
; FILING DATE: 12-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-55556-1/BIR;PROP-08-1
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 9 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-972-032-15

Query Match          25.0%; Score 5; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
DB      5 PRGAP 9

RESULT 3
US-08-642-255-15
; Sequence 15, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A55556-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEFAX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-642-255-15

Query Match          25.0%; Score 5; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
DB      5 PRGAP 9

RESULT 4
US-09-124-671-6
; Sequence 6, Application US/09124671A
; Patent No. 6160088
; GENERAL INFORMATION:
; APPLICANT: Rothman, James
```


; APPLICANT: Mayhew, Marx
; APPLICANT: Hoe, Mee
; TITLE OF INVENTION: KDEL RECEPTOR INHIBITORS
; FILE REFERENCE: 31488
; CURRENT APPLICATION NUMBER: US/09/124,671A
; CURRENT FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-124-671-6

Query Match 25.0%; Score 5; DB 3; Length 46;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETMFL 20
|||
Db 31 ETMFL 35

RESULT 5

US-09-134-001C-3074
; Sequence 3074, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3074
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3074

Query Match 25.0%; Score 5; DB 3; Length 74;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||
Db 58 PRGAP 62

RESULT 6

US-09-134-001C-4851
; Sequence 4851, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4851
; LENGTH: 74
; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4851

Query Match 25.0%; Score 5; DB 3; Length 74;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||
Db 58 PRGAP 62

RESULT 7

US-09-270-767-36912
; Sequence 36912, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36912
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-36912

Query Match 25.0%; Score 5; DB 4; Length 77;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||
Db 70 PRGAP 74

RESULT 8

US-09-270-767-52129
; Sequence 52129, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52129
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-52129

Query Match 25.0%; Score 5; DB 4; Length 77;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||
Db 70 PRGAP 74

RESULT 9

US-09-461-697-464
; Sequence 464, Application US/09461697

```
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranbam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 464
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-697-464

Query Match      25.0%; Score 5; DB 3; Length 80;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PRGAP 5
Db      54 PRGAP 58
      |||||

RESULT 10
US-09-252-991A-24182
; Sequence 24182, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24182
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24182

Query Match      25.0%; Score 5; DB 4; Length 91;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 RGAPM 6
Db      37 RGAPM 41
      |||||

RESULT 11
US-09-270-767-38667
; Sequence 38667, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
```

```
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38667
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-38667

Query Match      25.0%; Score 5; DB 4; Length 91;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PRGAP 5
Db      65 PRGAP 69
      |||||

RESULT 12
US-09-270-767-53884
; Sequence 53884, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53884
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-53884

Query Match      25.0%; Score 5; DB 4; Length 91;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PRGAP 5
Db      65 PRGAP 69
      |||||

RESULT 13
US-09-489-039A-13486
; Sequence 13486, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13486
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13486

Query Match      25.0%; Score 5; DB 4; Length 105;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      11 VCQML 15
Db      21 VCQML 25
      |||||
```

```
RESULT 14
US-09-124-671-21
; Sequence 21, Application US/09124671A
; Patent No. 6160088
; GENERAL INFORMATION:
; APPLICANT: Rothman, James
; APPLICANT: Mayhew, Mark
; APPLICANT: Hoe, Mee
; TITLE OF INVENTION: KDEL RECEPTOR INHIBITORS
; FILE REFERENCE: 31488
; CURRENT APPLICATION NUMBER: US/09/124,671A
; CURRENT FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric Xenopus laevis TSP4-KDEL
US-09-124-671-21

Query Match      25.0%; Score 5; DB 3; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      16 ETMFL 20
Db      60 ETMFL 64

RESULT 15
US-09-489-039A-11951
; Sequence 11951, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11951
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11951

Query Match      25.0%; Score 5; DB 4; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
Db      93 PRGAP 97

RESULT 16
US-09-270-767-34903
; Sequence 34903, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34903

Query Match      25.0%; Score 5; DB 4; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
Db      50 PRGAP 54

RESULT 17
US-09-270-767-50120
; Sequence 50120, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50120
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-50120

Query Match      25.0%; Score 5; DB 4; Length 125;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
Db      50 PRGAP 54

RESULT 18
US-09-252-991A-30594
; Sequence 30594, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30594
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30594

Query Match      25.0%; Score 5; DB 4; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
Db      50 PRGAP 54
```

```
Db          47 PRGAP 51
;
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23275
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23275

Query Match          25.0%; Score 5; DB 4; Length 160;
Best Local Similarity 100.0%; Pred.No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1 PRGAP 5
          |||||
Db          23 PRGAP 27

RESULT 22
US-09-134-000C-3710
; Sequence 3710, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3710
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3710

Query Match          25.0%; Score 5; DB 4; Length 163;
Best Local Similarity 100.0%; Pred.No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          2 RGAPM 6
          |||||
Db          12 RGAPM 16

RESULT 23
US-09-489-039A-8944
; Sequence 8944, Application US/09489039A
; Patent No. 6610816
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8944
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8944

Query Match          25.0%; Score 5; DB 4; Length 169;
Best Local Similarity 100.0%; Pred.No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db          57 PRGAP 61

RESULT 20
US-09-270-767-54985
; Sequence 54985, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54985
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54985

Query Match          25.0%; Score 5; DB 4; Length 139;
Best Local Similarity 100.0%; Pred.No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1 PRGAP 5
          |||||
Db          57 PRGAP 61

RESULT 21
US-09-270-767-54985
; Sequence 54985, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54985
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54985

Query Match          25.0%; Score 5; DB 4; Length 139;
Best Local Similarity 100.0%; Pred.No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1 PRGAP 5
          |||||
Db          57 PRGAP 61

RESULT 21
US-09-252-991A-23275
; Sequence 23275, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
```

QY 3 GAPMW 7
 |||||
 DB 90 GAPMW 94

RESULT 24
 US-09-252-991A-21850
 ; Sequence 21850, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 21850
 ; LENGTH: 183
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-21850

Query Match 25.0%; Score 5; DB 4; Length 183;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 |||||
 DB 107 PRGAP 111

RESULT 25
 US-09-252-991A-32100
 ; Sequence 32100, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 32100
 ; LENGTH: 185
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-32100

Query Match 25.0%; Score 5; DB 4; Length 185;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 |||||
 DB 170 PRGAP 174

RESULT 26
 US-09-270-767-47191
 ; Sequence 47191, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 47191
 ; LENGTH: 199
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 US-09-270-767-47191

Query Match 25.0%; Score 5; DB 4; Length 199;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
 |||||
 DB 109 QMLET 113

RESULT 27
 US-09-252-991A-26521
 ; Sequence 26521, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 26521
 ; LENGTH: 212
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-26521

Query Match 25.0%; Score 5; DB 4; Length 212;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 RWVCQ 13
 |||||
 DB 34 RWVCQ 38

RESULT 28
 US-09-270-767-60336
 ; Sequence 60336, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 60336
 ; LENGTH: 220
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 ; FEATURE:
 ; OTHER INFORMATION: Xaa means any amino acid
 US-09-270-767-60336

Query Match 25.0%; Score 5; DB 4; Length 220;

```
; Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
      |||||
Db      90 PRGAP 94

RESULT 29
US-09-252-991A-23304
; Sequence 23304, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23304
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23304

Query Match      25.0%; Score 5; DB 4; Length 237;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
      |||||
Db      40 PRGAP 44

RESULT 30
US-09-134-001C-4165
; Sequence 4165, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4165
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4165

Query Match      25.0%; Score 5; DB 3; Length 239;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      11 VCQML 15
      |||||
Db      52 VCQML 56

RESULT 31
US-09-252-991A-28343
; Sequence 28343, Application US/09252991A
```

```
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28343
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28343

Query Match      25.0%; Score 5; DB 4; Length 242;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
      |||||
Db      23 PRGAP 27

RESULT 32
US-08-426-819A-37
; Sequence 37, Application US/08426819A
; Patent No. 5723318
; GENERAL INFORMATION:
; APPLICANT: Yamaguchi, No. 5723318omi
; APPLICANT: Kojima, Tetsuo
; APPLICANT: Oh-Eda, Masayoshi
; APPLICANT: Hattori, Kunihiro
; TITLE OF INVENTION: Genes Coding for Megakaryocyte
; TITLE OF INVENTION: Potentiator
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/426,819A
; APPLICATION NUMBER: US/08/426,819A
; FILING DATE: 21-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 230-107P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
US-08-426-819A-37
```

Query Match 25.0%; Score 5; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 2.3e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||
Db 115 PRGAP 119

RESULT 33

US-09-248-796A-27518
; Sequence 27518, Application US/09248796A
; Patent No. 6747137

; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 27518
; LENGTH: 257

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-27518

Query Match 25.0%; Score 5; DB 4; Length 257;
Best Local Similarity 100.0%; Pred. No. 2.3e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
|||
Db 180 QMLET 184

RESULT 34

US-09-252-991A-21547
; Sequence 21547, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 21547
; LENGTH: 259

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-21547

Query Match 25.0%; Score 5; DB 4; Length 259;
Best Local Similarity 100.0%; Pred. No. 2.3e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||
Db 56 PRGAP 60

RESULT 35

Query Match 25.0%; Score 5; DB 1; Length 279;

US-09-710-279-1482
; Sequence 1482, Application US/09710279
; Patent No. 6703492

; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN

; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279

; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1482
; LENGTH: 276

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence

US-09-710-279-1482

Query Match 25.0%; Score 5; DB 4; Length 276;
Best Local Similarity 100.0%; Pred. No. 2.5e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCQML 15
|||
Db 89 VCQML 93

RESULT 36

US-08-312-387B-8

; Sequence 8, Application US/08312387B

; Patent No. 5545553

; GENERAL INFORMATION:

; APPLICANT: Gotschlich, Emil C.

; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM

; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/312,387B

; FILING DATE: July 7, 1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.

; REGISTRATION NUMBER: 26,742

; REFERENCE/DOCKET NUMBER: 600-1-095

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201 487-5800

; TELEFAX: 201 343-1684

; TELEX: 133521

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 279 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-312-387B-8

Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
|||||
Db 121 LETMF 125

RESULT 37
US-08-683-426-8
; Sequence 8, Application US/08683426
; Patent No. 5705367
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,426
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: September 26, 1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-683-426-8

Query Match 25.0%; Score 5; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
|||||
Db 121 LETMF 125

RESULT 38
US-08-683-458-8
; Sequence 8, Application US/08683458
; Patent No. 5798233
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson

; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,458
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: September 26, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-683-458-8

Query Match 25.0%; Score 5; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
|||||
Db 121 LETMF 125

RESULT 39
US-08-878-360-8
; Sequence 8, Application US/08878360
; Patent No. 5945322
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,360
; FILING DATE: 18-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/683,426
; FILING DATE:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: September 26, 1994


```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-878-360-8
;
Query Match 25.0%; Score 5; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
Db 121 LETMF 125

RESULT 40
US-09-333-412-8
; Sequence 8, Application US/09333412
; Patent No. 6342382
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; OLIGOSACCHARIDES, AND GENES ENCODING THEM
;
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/333,412
; FILING DATE: 15-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: July 7, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-09-333-412-8
;
Query Match 25.0%; Score 5; DB 3; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
```

```
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
Db 121 LETMF 125

RESULT 41
US-10-007-267-8
; Sequence 8, Application US/10007267
; Patent No. 6780624
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; OLIGOSACCHARIDES, AND GENES ENCODING THEM
;
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/007,267
; FILING DATE: 03-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/333,412
; FILING DATE: 15-Jun-1999
; APPLICATION NUMBER: 08/312,387
; FILING DATE: July 7, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-10-007-267-8
;
Query Match 25.0%; Score 5; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
Db 121 LETMF 125

RESULT 42
US-08-312-387B-6
; Sequence 6, Application US/08312387B
; Patent No. 5545553
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; OLIGOSACCHARIDES, AND GENES ENCODING THEM
;
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/312,387B
FILING DATE: July 7, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-312-387B-6

Query Match 25.0%; Score 5; DB 1; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.5e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
DB 121 LETMF 125

RESULT 43
US-08-683-426-6
Sequence 6, Application US/08683426
Patent No. 5705367
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,426
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095B

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-426-6
Query Match 25.0%; Score 5; DB 1; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.5e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 LETMF 19
DB 121 LETMF 125
RESULT 44
US-08-683-458-6
Sequence 6, Application US/08683458
Patent No. 5798233
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,458
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-458-6

Query Match 25.0%; Score 5; DB 1; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.5e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
DB 121 LETMF 125

```
RESULT 45
US-08-878-360-6
; Sequence 6, Application US/08878360
; Patent No. 5945322
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,360
; FILING DATE: 18-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 08/683,426
; FILING DATE:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: September 26, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-878-360-6
Query Match 25.0%; Score 5; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
Db 121 LETMF 125

RESULT 46
US-08-478-140B-6
; Sequence 6, Application US/08478140B
; Patent No. 6127153
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, KARL F.
; APPLICANT: ROTH, STEPHEN
; APPLICANT: BUCZALA, STEPHANIE L.
; TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO
; TITLE OF INVENTION: SACCHARIDE UNITS WITH A POLYGLYCOSYLTRANSFERASE, A
; TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE AND GENE ENCODING A
; TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
```

```
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,140B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Laura A. Coruzzi
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7188-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-478-140B-6
Query Match 25.0%; Score 5; DB 3; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
Db 121 LETMF 125

RESULT 47
US-09-333-412-6
; Sequence 6, Application US/09333412
; Patent No. 6342382
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/333,412
; FILING DATE: 15-JUN-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: July 7, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
```

TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-333-412-6

Query Match 25.0%; Score 5; DB 3; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
DB 121 LETMF 125

RESULT 48
US-09-338-943-6
Sequence 6, Application US/09338943
Patent No. 6379933
GENERAL INFORMATION:
APPLICANT: JOHNSON, KARL F.
APPLICANT: ROTH, STEPHEN
APPLICANT: BUCZALA, STEPHANIE L.
TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO
TITLE OF INVENTION: SACCHARIDE UNITS WITH A POLYGLYCOSYLTRANSFERASE, A
TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE AND GENE ENCODING A
TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/338,943
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,140
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Laura A. Coruzzi
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7188-017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-338-943-6

Query Match 25.0%; Score 5; DB 3; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19

DB 121 LETMF 125
RESULT 49
US-10-007-267-6
Sequence 6, Application US/10007267
Patent No. 6780624
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,267
FILING DATE: 03-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/333,412
FILING DATE: 15-Jun-1999
APPLICATION NUMBER: 08/312,387
FILING DATE: July 7, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-007-267-6

Query Match 25.0%; Score 5; DB 4; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETMF 19
DB 121 LETMF 125

RESULT 50
US-09-904-615-106
Sequence 106, Application US/09904615
Patent No. 6566325
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 49 Human Secreted Proteins
FILE REFERENCE: P2032P1
CURRENT APPLICATION NUMBER: US/09/904,615
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 09/511,554
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/097,917

us-10-066-965a-3.oligo.ra1

Tue Oct 26 08:02:20 2004

```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25496

Query Match      25.0%; Score 5; DB 4; Length 289;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
      |||||
Db      236 PRGAP 240

RESULT 53
US-09-328-352-7849
; Sequence 7849, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7849
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7849

Query Match      25.0%; Score 5; DB 4; Length 297;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      15 LETMF 19
      |||||
Db      172 LETMF 176

RESULT 54
US-09-583-110-5101
; Sequence 5101, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5101
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-5101

Query Match      25.0%; Score 5; DB 4; Length 304;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      14 MLETM 18
      |||||
Db      111 MLETM 115
```

```
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 106
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (283)
; OTHER INFORMATION: Xaa equals stop translation
US-09-904-615-106

Query Match      25.0%; Score 5; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
      |||||
Db      199 PRGAP 203

RESULT 51
US-09-252-991A-27676
; Sequence 27676, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27676
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27676

Query Match      25.0%; Score 5; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
      |||||
Db      96 PRGAP 100

RESULT 52
US-09-252-991A-25496
; Sequence 25496, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25496
; LENGTH: 289
```

```
RESULT 55
US-08-789-609A-2
; Sequence 2, Application US/08789609A
; Patent No. 5827689
; GENERAL INFORMATION:
; APPLICANT: Gentry, Daniel
; APPLICANT: Lonsdale, John
; APPLICANT: Pearson, Stewart
; APPLICANT: Payne, David
; TITLE OF INVENTION: NO. 5827689e1 FabD
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; APPLICATION NUMBER: US/08/789,609A
; FILING DATE: 24-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030685
; FILING DATE: 13-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: GM50004
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 308 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-789-609A-2

Query Match 25.0%; Score 5; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LETMF 19
Db 45 LETMF 49

RESULT 56
US-09-108-517-2
; Sequence 2, Application US/09108517
; Patent No. 6258934
; GENERAL INFORMATION:
; APPLICANT: Gentry, Daniel
; APPLICANT: Lonsdale, John
; APPLICANT: Pearson, Stewart
; APPLICANT: Payne, David
; TITLE OF INVENTION: No. 6258934e1 FabD
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
```

```
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/108,517
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/789,609
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: GM50004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 308 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-108-517-2

Query Match 25.0%; Score 5; DB 3; Length 308;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LETMF 19
Db 45 LETMF 49

RESULT 57
US-10-138-701-16
; Sequence 16, Application US/10138701
; Patent No. 6753149
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc. et al.
; TITLE OF INVENTION: Staphylococcus aureus genes and polypeptides
; FILE REFERENCE: PB484
; CURRENT APPLICATION NUMBER: US/10/138,701
; CURRENT FILING DATE: 2002-05-06
; PRIOR FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US/09/512,255A
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/098,964
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-05
; PRIOR APPLICATION NUMBER: PCT/ US99/19726
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-138-701-16

Query Match 25.0%; Score 5; DB 4; Length 308;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LETMF 19
Db 45 LETMF 49
```

```
Db          45 LETMF 49

RESULT 58
US-09-134-001C-4824
; Sequence 4824, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4824
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4824

Query Match          25.0%; Score 5; DB 3; Length 311;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          15 LETMF 19
           |||||
Db          48 LETMF 52

RESULT 59
US-09-252-991A-19219
; Sequence 19219, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19219
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19219

Query Match          25.0%; Score 5; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 PRGAP 5
           |||||
Db          106 PRGAP 110

RESULT 60
US-09-252-991A-19369
; Sequence 19369, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19369
; LENGTH: 339
; TYPE: PRT

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19369
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19369

Query Match          25.0%; Score 5; DB 4; Length 327;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 PRGAP 5
           |||||
Db          141 PRGAP 145

RESULT 61
US-09-252-991A-32186
; Sequence 32186, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32186
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32186

Query Match          25.0%; Score 5; DB 4; Length 330;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          9 RWVCQ 13
           |||||
Db          200 RWVCQ 204

RESULT 62
US-09-252-991A-19891
; Sequence 19891, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19891
; LENGTH: 339
; TYPE: PRT
```

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19891

Query Match 25.0%; Score 5; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 94 PRGAP 98

RESULT 63

US-09-270-767-46634
; Sequence 46634, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46634
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-46634

Query Match 25.0%; Score 5; DB 4; Length 354;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMMW 8
|||||
Db 165 APMMW 169

RESULT 64

US-09-252-991A-20554
; Sequence 20554, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20554
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (168), (202), (214), (251)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-20554

Query Match 25.0%; Score 5; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||

Db 18 PRGAP 22

RESULT 65

US-09-252-991A-17858
; Sequence 17858, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,788
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17858
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17858

Query Match 25.0%; Score 5; DB 4; Length 363;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
|||||
Db 50 RGAPM 54

RESULT 66

US-09-252-991A-31139
; Sequence 31139, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31139
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31139

Query Match 25.0%; Score 5; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
Db 153 PRGAP 157

RESULT 67

US-09-252-991A-23047
; Sequence 23047, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; OTHER INFORMATION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23047
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23047

Query Match 25.0%; Score 5; DB 4; Length 378;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||
Db 200 PRGAP 204

RESULT 68
US-09-252-991A-19344
; Sequence 19344, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19344
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19344

Query Match 25.0%; Score 5; DB 4; Length 394;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||
Db 366 PRGAP 370

RESULT 69
US-09-252-991A-26116
; Sequence 26116, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26116
; LENGTH: 395
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26116

Query Match 25.0%; Score 5; DB 4; Length 395;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||
Db 384 PRGAP 388

RESULT 70
US-09-252-991A-30428
; Sequence 30428, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30428
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30428

Query Match 25.0%; Score 5; DB 4; Length 420;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||
Db 352 PRGAP 356

RESULT 71
US-09-328-352-5815
; Sequence 5815, Application US/09328352
; Patent No. 6582958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5815
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5815

Query Match 25.0%; Score 5; DB 4; Length 421;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLET 17
|||
Db 190 QMLET 194

RESULT 72
US-08-396-218-2
; Sequence 2, Application US/08396218

Patent No. 5695966
; GENERAL INFORMATION:
; APPLICANT: INVENTI, Augusto
; APPLICANT: BREME, Umberto
; APPLICANT: COLOMBO, Anna L
; APPLICANT: HUTCHINSON, Charles R
; APPLICANT: OTTEN, Sharee
; APPLICANT: SCOTTI, Claudio
; TITLE OF INVENTION: PROCESS FOR PREPARING DOXORUBICIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAI DO, MARMELSTEIN, MURRAY & ORAM
; STREET: 655 Fifteenth Street, N. W., Suite 330 - G
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/396,218
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KITTS, Monica C
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: P1615-5002
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-396-218-2

Query Match 25.0%; Score 5; DB 1; Length 422;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 325 PRGAP 329

RESULT 73
US-08-760-116-2
; Sequence 2, Application US/08760116
; Patent No. 5786190
; GENERAL INFORMATION:
; APPLICANT: INVENTI, Augusto
; APPLICANT: BREME, Umberto
; APPLICANT: COLOMBO, Anna L
; APPLICANT: HUTCHINSON, Charles R
; APPLICANT: OTTEN, Sharee
; APPLICANT: SCOTTI, Claudio
; TITLE OF INVENTION: PROCESS FOR PREPARING DOXORUBICIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAI DO, MARMELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W., Suite 330 - G
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,116
; FILING DATE: 3-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/396,218
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KITTS, Monica C
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: P1615-6007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-760-116-2

Query Match 25.0%; Score 5; DB 1; Length 422;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 325 PRGAP 329

RESULT 74
US-09-096-982-5
; Sequence 5, Application US/09096982
; Patent No. 5962293
; GENERAL INFORMATION:
; APPLICANT: Strohl, William R.
; APPLICANT: Dickens, Michael L.
; APPLICANT: DeSanti, Charles L.
; TITLE OF INVENTION: METHOD OF PRODUCING DOXORUBICIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CALFEE, HALTER & GRISWOLD
; STREET: 800 Superior Avenue, Suite 1400
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: USA
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/096,982
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Goltick, Mary E.
; REGISTRATION NUMBER: 34829
; REFERENCE/DOCKET NUMBER: 22727/00131
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 216-622-8458
; TELEFAX: 216-241-0816
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-09-096-982-5

Query Match 25.0%; Score 5; DB 2; Length 422;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
DB 325 PRGAP 329

RESULT 75
US-08-653-650A-5
; Sequence 5, Application US/08653650A
; Patent No. 5976830
; GENERAL INFORMATION:
; APPLICANT: Strohl, William R.
; APPLICANT: Dickens, Michael L.
; APPLICANT: Desanti, Charles L.
; TITLE OF INVENTION: METHOD OF PRODUCING DOXORUBICIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CALFEE, HALTER & GRISWOLD
; STREET: 800 Superior Avenue, Suite 1400
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: USA
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/653,650A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Golrick, Mary E.
; REGISTRATION NUMBER: 34829
; REFERENCE/DOCKET NUMBER: 22727/00131
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 216-622-8458
; TELEFAX: 216-241-0816
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-653-650A-5

Query Match 25.0%; Score 5; DB 2; Length 422;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
DB 325 PRGAP 329

Search completed: October 26, 2004, 07:25:06
Job time : 24.25 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 07:05:46 ; Search time 16.5 Seconds
(without alignments)
116.626 Million cell updates/sec

Title: US-10-066-965A-4

Perfect score: 20

Sequence: 1 PRGAPWMLRCVCMLETKFL 20

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	30.0	238	2	S75336
2	6	30.0	278	2	D83815
3	6	30.0	421	2	A82744
4	6	30.0	434	2	G82372
5	6	30.0	767	2	T19690
6	6	30.0	3396	2	T22613
7	5	25.0	34	2	PQ0341
8	5	25.0	34	2	PQ0363
9	5	25.0	70	2	T18008
10	5	25.0	75	2	B54188
11	5	25.0	75	2	T26778
12	5	25.0	76	2	AB1212
13	5	25.0	77	2	F69219
14	5	25.0	90	2	G69133
15	5	25.0	102	2	A72518
16	5	25.0	113	2	AF2744
17	5	25.0	114	2	A55010
18	5	25.0	114	2	E97525
19	5	25.0	119	2	T08271
20	5	25.0	123	2	T11049
21	5	25.0	123	2	D72579
22	5	25.0	127	2	AI0329
23	5	25.0	127	2	AB0154
24	5	25.0	128	1	S12372
25	5	25.0	129	1	I53637
26	5	25.0	132	2	A57325
27	5	25.0	142	2	G82796
28	5	25.0	147	2	T50331
29	5	25.0	148	2	AH1645

30	5	25.0	148	2	AD1594
31	5	25.0	162	2	D75295
32	5	25.0	165	2	F82743
33	5	25.0	167	2	AH1106
34	5	25.0	173	2	H81294
35	5	25.0	179	2	S58124
36	5	25.0	192	2	D69061
37	5	25.0	192	2	T46483
38	5	25.0	193	2	AH1855
39	5	25.0	195	1	SAVLDM
40	5	25.0	195	1	SAVLDS
41	5	25.0	199	1	QQBEC5
42	5	25.0	202	2	I46688
43	5	25.0	202	2	A97342
44	5	25.0	206	2	D70335
45	5	25.0	218	2	S28710
46	5	25.0	222	2	S33204
47	5	25.0	229	2	B95910
48	5	25.0	229	2	B81429
49	5	25.0	233	2	B90222
50	5	25.0	239	2	T29697
51	5	25.0	240	2	T20657
52	5	25.0	247	2	B89833
53	5	25.0	248	1	LNDGFS
54	5	25.0	249	2	T05424
55	5	25.0	250	2	C70823
56	5	25.0	253	2	S49055
57	5	25.0	253	2	G90275
58	5	25.0	254	2	C98274
59	5	25.0	262	2	T29698
60	5	25.0	266	2	H83956
61	5	25.0	267	2	T00114
62	5	25.0	273	2	T21429
63	5	25.0	277	2	S16563
64	5	25.0	278	2	C69854
65	5	25.0	278	2	T27610
66	5	25.0	280	2	E70745
67	5	25.0	284	2	T25011
68	5	25.0	296	2	T24827
69	5	25.0	304	2	G72776
70	5	25.0	305	2	T20906
71	5	25.0	306	2	AD2202
72	5	25.0	308	2	H75049
73	5	25.0	308	2	G71110
74	5	25.0	317	2	T49591
75	5	25.0	318	2	T33845
76	5	25.0	325	2	AB2574
77	5	25.0	325	2	AH1312
78	5	25.0	325	2	AH1684
79	5	25.0	330	2	T02347
80	5	25.0	334	2	C82935
81	5	25.0	335	2	F86586
82	5	25.0	335	2	C72038
83	5	25.0	342	2	AF1398
84	5	25.0	343	1	C69211
85	5	25.0	348	2	AI1889
86	5	25.0	351	2	D96761
87	5	25.0	355	2	T29932
88	5	25.0	357	2	T24137
89	5	25.0	359	2	T13289
90	5	25.0	368	2	T40115
91	5	25.0	382	2	D82264
92	5	25.0	387	2	JE0364
93	5	25.0	388	2	T09489
94	5	25.0	389	1	A48329
95	5	25.0	390	2	A60093
96	5	25.0	393	2	S39383
97	5	25.0	406	2	S39965
98	5	25.0	407	2	T06693
99	5	25.0	409	2	T47298
100	5	25.0	409	2	T24138

hypothetical prote
DNA topology modul
hypothetical prote
hypothetical prote
probable lipoprote
hypothetical prote
precorrin-8W decar
hypothetical prote
hypothetical prote
delta large antige
complement compone
HQLF2 protein prec
nudix (MutT) famil
hypothetical prote
RNA-directed RNA p
hypothetical prote
conserved hypotet
hypothetical prote
conserved hypotet
hypothetical prote
hypothetical prote
pulmonary surfacta
hypothetical prote
thioesterase - Str
hypothetical prote
nikP2 protein (AJ2
flagellar assembly
exodeoxyribonuclea
hypothetical prote
molG protein - Rhl
conserved hypotet
hypothetical prote
hypothetical prote
hypothetical prote
transcription regu
transcription regu
probable lipid tra
recombination prot
oxidoreductase [im
conserved hypotet
ATP binding protei
conserved hypotet
hypothetical prote
hypothetical prote
probable integrase
uv excision repair
probable exopolysa
lactosylceramide a
hypothetical prote
histidinol-phospha
cytokeratin, type
cyclin CCL1 - Yeas
hypothetical prote
hypothetical prote
probable replicati
hypothetical prote

ALIGNMENTS

```

RESULT 1
S75336
hypothetical protein sll1925 - Synecocystis sp. (strain PCC 6803)
C:Species: Synecocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S75336
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75336
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-238 <KAN>
A:Cross-references: UNIPROT:P73223; EMBL:D90904; GB:AB001339; NID:g1652225; PIDN:BAAL1725
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 30.0%; Score 6; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKEL 20
|||||
DB 165 LETKEL 170

RESULT 2
D83815
shikimate 5-dehydrogenase aroD [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 16-Aug-2004
C:Accession: D83815
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: D83815
A>Status: preliminary
A:Cross-references: UNIPROT:Q9KD93; GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BA050
A:Experimental source: strain C-125
C:Genetics:
A:Gene: aroD
C:Superfamily: Shikimate 5-dehydrogenase; shikimate dehydrogenase homology

Query Match 30.0%; Score 6; DB 2; Length 278;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKEL 20
|||||
DB 225 LETKEL 230

RESULT 3
A82744
conserved hypothetical protein XF0928 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: A82744
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: A82744

```

```

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-421 <SIM>
A:Cross-references: UNIPROT:Q9PEV0; GB:AE003932; GB:AE003849; NID:g9105849; PIDN:AAF8373
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A:Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0928
C:Superfamily: hypothetical protein HI0624

Query Match 30.0%; Score 6; DB 2; Length 421;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMWLR 9
|||||
DB 159 APMWLR 164

RESULT 4
G82372
sun protein VC0044 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82372
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82372
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-434 <HEI>
A:Cross-references: UNIPROT:Q9KVU5; GB:AE004096; GB:AE003852; NID:g9654440; PIDN:AAF9332
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0044
A:Map position: 1
C:Superfamily: hypothetical protein HI0624

Query Match 30.0%; Score 6; DB 2; Length 434;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMWLR 9
|||||
DB 177 APMWLR 182

RESULT 5
T19690
hypothetical protein C33G3.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T19690
R:Matthews, L.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19163

```

```

A:Accession: T19690
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-767 <WIL>
A:Cross-references: UNIPROT:Q33325; EMBL:Z78540; PIDN:CA801738.1; GSPDB:GN000028; CBSP:CF
A:Experimental source: clone C33G3
C:Genetics:
A:Gene: CBSP:C33G3.1
A:Map position: X
A:Introns: 57/3; 72/3; 95/1; 143/3; 240/3; 285/3; 336/3; 419/3; 585/2; 658/3; 690/3; 757/3
Query Match 30.0%; Score 6; DB 2; Length 767;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 QMLETK 18
Db 403 QMLETK 408

RESULT 6
T22613
hypothetical protein F54B3.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22613
R:McMurray, A.
submitted to the EMBL Data Library, March 1995
A:Reference number: Z19589
A:Accession: T22613
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-3396 <WIL>
A:Cross-references: UNIPROT:O62263; EMBL:Z48583; PIDN:CAA8472.1; GSPDB:GN000020; CBSP:FS
A:Experimental source: clone F54B3
C:Genetics:
A:Gene: CBSP:F54B3.1
A:Map position: 2
A:Introns: 52/3; 109/3; 269/1; 432/1; 560/2; 747/2; 938/3; 1073/1; 1719/2; 1877/2
C:Superfamily: Caenorhabditis elegans hypothetical protein F54B3.1

Query Match 30.0%; Score 6; DB 2; Length 3396;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 QMLETK 18
Db 1962 QMLETK 1967

RESULT 7
PQ0341
L protein - rabies virus
C:Species: rabies virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PQ0341
R:Sacramento, D.; Badrane, H.; Bourhy, H.; Tordo, N.
J. Gen. Virol. 73, 1149-1158, 1992
A:Title: Molecular epidemiology of rabies virus in France: Comparison with vaccine strain
A:Reference number: PQ0340; MUID:92268873; PMID:1588319
A:Accession: PQ0341
A:Molecule type: mRNA
A:Residues: 1-34 <SAC>
A:Cross-references: UNIPROT:Q9IPJ5
A>Note: this sequence was obtained from isolates: WR56, WR06, WR39, WR97, WR18, WR63, WR
C:Superfamily: rhabdovirus L protein

Query Match 25.0%; Score 5; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PRGAP 5
Db 403 QMLETK 408

```

```

Db 21 PRGAP 25

RESULT 8
PQ0363
L protein - rabies virus (isolate WR17)
C:Species: rabies virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PQ0363
R:Sacramento, D.; Badrane, H.; Bourhy, H.; Tordo, N.
J. Gen. Virol. 73, 1149-1158, 1992
A:Title: Molecular epidemiology of rabies virus in France: Comparison with vaccine strain
A:Reference number: PQ0340; MUID:92268873; PMID:1588319
A:Accession: PQ0363
A:Molecule type: mRNA
A:Residues: 1-34 <SAC>
A:Cross-references: UNIPROT:Q9IPJ5
C:Superfamily: rhabdovirus L protein

Query Match 25.0%; Score 5; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 9
T18008
hypothetical protein a506R - Chlorella virus PBCV-1
C:Species: Chlorella virus PBCV-1
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18008
R:Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z18806
A:Accession: T18008
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-70 <GRA>
A:Cross-references: UNIPROT:Q98556; EMBL:U42580; MID:G4028996; PIDN:AAC96873.1
A:Experimental source: specific host Chlorella strain NC64A
C:Genetics:
A>Note: a506R

Query Match 25.0%; Score 5; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 LRCVC 12
Db 54 LRCVC 58

RESULT 10
B54188
granulocyte chemotactic protein, GCP-2 - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 13-Sep-1994 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
C:Accession: B54188
R:Proost, P.; Wuyts, A.; Conings, R.; Lenaerts, J.P.; Billiau, A.; Opendakker, G.; Van De
Biochemistry 32, 10170-10177, 1993
A:Title: Human and bovine granulocyte chemotactic protein-2: complete amino acid sequence
A:Reference number: A54188; MUID:94001982; PMID:8399143
A:Accession: B54188
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-75 <PRO>
A:Experimental source: MDBK cells
A>Note: sequence extracted from NCBI backbone (NCBIP:137967)
C:Superfamily: beta-thromboglobulin

```

```
Query Match      25.0%; Score 5; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 LRCVC 12
      |||||
DB     10 LRCVC 14

RESULT 11
T26778
Hypothetical protein Y40B1A.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26778
R:Harris, B.
submitted to the EMBL Data Library, December 1998
A:Reference number: Z20264
A:Accession: T26778
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-75 <WIL>
A:Cross-references: UNIPROT:Q9XH29; EMBL:AL034392; PIDN:CAA22304.1; CESP:Y40B1A.1
A:Experimental source: clone Y40B1A
C:Genetics:
A:Gene: CESP:Y40B1A.1
A:Introns: 61/3

Query Match      25.0%; Score 5; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
      |||||
DB      5 PRGAP 9

RESULT 12
AB1212
TN916 ORF8 homolog lmo1098 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB1212
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1212
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-76 <GLA>
A:Cross-references: UNIPROT:Q8Y020; GB:NC_003210; PIDN:CAC99176.1; PID:gl6410500; GSPDB:
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo1098

Query Match      25.0%; Score 5; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      14 MLETK 18
      |||||
DB      61 MLETK 65

RESULT 13
F69219
conserved hypothetical protein MTH895 - Methanobacterium thermoautotrophicum (strain Del
C:Species: Methanobacterium thermoautotrophicum
```

```
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: F69219
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.
; Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: F69219
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-77 <MTH>
A:Cross-references: UNIPROT:O26981; GB:AE000665; GB:AE000666; NID:g2621984; PIDN:AAB8539
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH895
C:Superfamily: probable glutaredoxin grx-2

Query Match      25.0%; Score 5; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      12 COMLE 16
      |||||
DB      14 COMLE 18

RESULT 14
S69133
platelet factor 4 - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 10-Mar-1998 #sequence_revision 10-Mar-1998 #text_change 09-Jul-2004
C:Accession: S69133
R:Proudfoot, A.E.I.; Magnenat, E.; Haley, T.M.; Maione, T.E.; Wells, T.N.C.
Eur. J. Biochem. 228, 658-664, 1995
A:Title: The complete primary structure of glycosylated porcine platelet factor 4.
A:Reference number: S69133; MUID:95255268; PMID:7737160
A:Accession: S69133
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-90 <PRO>
A:Cross-references: UNIPROT:P30034
A:Note: blocked N-terminus
C:Superfamily: beta-thromboglobulin

Query Match      25.0%; Score 5; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 LRCVC 12
      |||||
DB      23 LRCVC 27

RESULT 15
A72518
Hypothetical protein APE2118 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: A72518
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: A72518
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-102 <KAW>
A:Cross-references: UNIPROT:Q9YA21; DDBJ:AP000063; NID:g5105654; PIDN:BAA81129.1; PID:g95
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE2118
```



```

C;Superfamily: Aeropyrum pernix hypothetical protein APE2118
  Query Match      25.0%; Score 5; DB 2; Length 102;
  Best Local Similarity 100.0%; Pred. No. 1.1e+02;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 4 APWWL 8
  Db 35 APWWL 39

RESULT 16
AF2744
hypothetical protein Atul366 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AF2744
R;Wood, D.W.; Secubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H.;
  erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutuyavin, T.; Levy, R.; Li, M.; McClell
  ; Karp, P.; Romero, P.; Zhang, S.
  Science 294, 2317-2323, 2001
A;Title: Cloning and characterization of the human neutrophil-activating peptide ENA-78
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
  ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AF2744
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <KUR>
A;Cross-references: UNIPROT:Q8UFN1; GB:AE008688; PIDN:AAL42372.1; PID:g17739780; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atul366
A;Map position: circular chromosome

  Query Match      25.0%; Score 5; DB 2; Length 113;
  Best Local Similarity 100.0%; Pred. No. 1.2e+02;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 5 PMWLR 9
  Db 71 PMWLR 75

RESULT 17
A55010
neutrophil-activating peptide ENA-78 precursor - human
N;Alternate names: epithelial-derived neutrophil-activating peptide 78 (ENA-78)
C;Species: Homo sapiens (man)
C;Date: 11-Nov-1994 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C;Accession: JC2433; A55010; I37357; JH0558; P50347; S44075
R;Corbett, M.S.; Schmitt, I.; Riess, O.; Walz, A.
  Biochem. Biophys. Res. Commun. 205, 612-617, 1994
A;Title: Characterization of the gene for human neutrophil-activating peptide 78 (ENA-78)
A;Reference number: JC2433; MUID:95091791; PMID:7999089
A;Accession: JC2433
A;Molecule type: DNA
A;Residues: 1-114 <COR>
A;Cross-references: UNIPROT:P42830; GB:I37036; NID:G607030; PIDN:AAA86426.1; PID:G607031
R;Chang, M.; McMinch, J.; Basu, R.; Simonet, S.
  J. Biol. Chem. 269, 25277-25282, 1994
A;Title: Cloning and characterization of the human neutrophil-activating peptide (ENA-78)
A;Reference number: A55010; MUID:95014315; PMID:7929219
A;Accession: A55010
A;Molecule type: DNA
A;Residues: 1-12, 'S', 14-114 <CHA>
R;Power, C.A.; Furness, R.B.; Brawand, C.; Wells, T.N.
  Gene 151, 333-334, 1994
A;Title: Cloning of a full-length cDNA encoding the neutrophil-activating peptide ENA-78
A;Reference number: I37357; MUID:95129887; PMID:7828901
A;Accession: I37357
A;Status: translated from GB/EMBL/DBJ

```

```

A;Molecule type: mRNA
A;Residues: 1-114 <RES>
A;Cross-references: EMBL:X78686; NID:g471242; PIDN:CAA55355.1; PID:g471243
R;Walz, A.; Burgener, R.; Car, B.; Baggiolini, M.; Kunkel, S.L.; Strieter, R.M.
  J. Exp. Med. 174, 1355-1362, 1991
A;Title: Structure and neutrophil-activating properties of a novel inflammatory peptide
A;Reference number: JH0558; MUID:92078844; PMID:1744577
A;Accession: JH0558
A;Molecule type: mRNA
A;Residues: 43-114 <WALI>
A;Experimental source: pulmonary type II epithelial cell line A549
A;Accession: PS0347
A;Molecule type: protein
A;Residues: 37-70;93-114 <WAL2>
C;Genetics:
A;Gene: ENA78; NAP
A;Map position: 4q13-q21
A;Introns: 37/1; 81/3; 109/2
C;Superfamily: beta-thromboglobulin
C;Keywords: cytokine
F;1-17/Domain: signal sequence #status predicted <SIG>
F;37-114/Product: neutrophil-activating peptide ENA-78 #status experimental <MAT>

  Query Match      25.0%; Score 5; DB 2; Length 114;
  Best Local Similarity 100.0%; Pred. No. 1.2e+02;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 8 LRCVC 12
  Db 47 LRCVC 51

RESULT 18
E97525
hypothetical protein AGR_C_2526 [imported] - Agrobacterium tumefaciens (strain C58, Cerec)
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: E97525
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
  A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
  Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: E97525
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-114 <KUR>
A;Cross-references: UNIPROT:Q8UFN1; GB:AE007869; PIDN:AAK87158.1; PID:g15156430; GSPDB:GN
C;Genetics:
A;Gene: AGR_C_2526
A;Map position: circular chromosome

  Query Match      25.0%; Score 5; DB 2; Length 114;
  Best Local Similarity 100.0%; Pred. No. 1.2e+02;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 5 PMWLR 9
  Db 72 PMWLR 76

RESULT 19
T08271
probable thioredoxin - Halobacterium sp. (strain NRC-1) plasmid pNRC100
N;Alternate names: hypothetical protein H0606; hypothetical protein H1757
C;Species: Halobacterium sp.
A;Variety: strain NRC-1
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T08271; T08372
R;Ng, W.V.; Ciufu, S.A.; Smith, T.M.; Bumgarner, R.E.; Baskin, D.; Faust, J.; Hall, B.; I
  Genome Res. 8, 1131-1141, 1998
A;Title: Snapshot of a large dynamic replicon in a halophilic Archaeon: megaplasmid or mi
A;Reference number: Z16408; MUID:99063795; PMID:9847077

```

A:Accession: T08271
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-119 <NGW>
 A:Cross-references: UNIPROT:O46709; EMBL:AF016485; NID:g28222278; PID:g28222332; HALOSP:HD
 A:Experimental source: strain NRC-1
 C:Genetics: COP1
 A:Accession: T08372
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-119 <DAS>
 A:Cross-references: EMBL:AF016485; NID:g2822278; PID:g2822433; HALOSP:HI757
 A:Experimental source: strain NRC-1
 C:Genetics: COP2
 A:Gene: trxA; HALOSP:H0606
 A:Genome: plasmid pNRC100
 C:Genetics: <COP2>
 A:Gene: trxA; HALOSP:HI757
 A:Genome: plasmid pNRC100
 C:Function:
 A:Description: involved in reduction of ribonucleotides, methionine sulfoxide sulfate, a
 C:Superfamily: thioredoxin, thioredoxin homology

 Query Match 25.0%; Score 5; DB 2; Length 119;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 12 COMLE 16
 |||||
 Db 48 COMLE 52

 RESULT 20
 T11049
 hypothetical protein ORF119 - Chlamydomonas eugametos mitochondrion
 C:Species: mitochondrion Chlamydomonas eugametos
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C:Accession: T11049
 R:Denovan-Wright, E.M.; Nedelcu, A.M.; Lee, R.W.
 Plant Mol. Biol. 36, 285-295, 1998
 A:Title: Complete sequence of the mitochondrial DNA of Chlamydomonas eugametos.
 A:Reference number: Z17244; MUID:98145434; PMID:9484440
 A:Accession: T11049
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-119 <DEN>
 C:Genetics:
 A:Cross-references: UNIPROT:O47553; EMBL:AF008237; NID:g2865253; PID:g2865265; PIDN:AAC3
 C:Genome: mitochondrion
 C:Keywords: mitochondrion

 Query Match 25.0%; Score 5; DB 2; Length 119;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 15 LETKF 19
 |||||
 Db 72 LETKF 76

 RESULT 21
 D72579
 hypothetical protein APE1916 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 C:Accession: D72579
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
 DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: D72579

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-123 <KAW>
 A:Cross-references: UNIPROT:Q9YAM7; DDBJ:AP000062; NID:g5105244; PIDN:BAA80921.1; PID:d1
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE1916

 Query Match 25.0%; Score 5; DB 2; Length 123;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 PRGAP 5
 |||||
 Db 34 PRGAP 38

 RESULT 22
 AI0329
 conserved hypothetical protein YPO2705 [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 16-Aug-2004
 C:Accession: AI0329
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
 geno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AI0329
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-127 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC92944.1; PID:gi5980683; GSPDB:GN00175
 C:Genetics:
 A:Superfamily: Glycyl radical cofactor protein YfiD; glycyl radical homology

 Query Match 25.0%; Score 5; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 8 LRCVC 12
 |||||
 Db 29 LRCVC 33

 RESULT 23
 AB0154
 probable bacteriophage protein (partial) YPO1259 [imported] - Yersinia pestis (strain CO
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
 C:Accession: AB0154
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
 geno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AB0154
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-127 <KUR>
 A:Cross-references: UNIPROT:Q8ZGM8; GB:AL590842; PIDN:CAC90093.1; PID:gi5979313; GSPDB:G
 C:Genetics:
 A:Gene: YPO1259

 Query Match 25.0%; Score 5; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 6 MWLRC 10
 |||||

Db 1 MWLRC 5

RESULT 24

S12372
 spasmolytic protein precursor - mouse (fragment)
 C/Species: Mus musculus (house mouse)
 C/Date: 21-Nov-1993 #sequence_revision 24-May-1996 #text_change 18-Jun-1999
 C/Accession: S12372; S63997
 R/Tomasetto, C.; Rio, M.C.; Gautier, C.; Wolf, C.; Hareuveni, M.; Chambon, P.; Lathe, R.
 EMBO J. 9, 407-414, 1990
 A/Title: hSP, the domain-duplicated homolog of pS2 protein, is co-expressed with pS2 in
 A/Reference number: S12371; MUID:90151615; PMID:2303034
 A/Accession: S12372
 A/Molecule type: mRNA
 A/Residues: 1-128 <TOM>
 A/Cross-references: EMBL:X51697
 R/Tomasetto, C.; Rio, M.C.; Gautier, C.; Wolf, C.; Hareuveni, M.; Chambon, P.; Lathe, R.
 submitted to the EMBL Data Library, December 1993
 A/Description: hSP, the domain-duplicated homolog of pS2 protein, is co-expressed with p
 A/Reference number: S63997
 A/Accession: S63997
 A/Molecule type: mRNA
 A/Residues: 1-10, V', 12-128 <TOW>
 A/Cross-references: EMBL:X51697; NID:G54165; PIDN:CAA35994.1; PID:G54166
 C/Genetics:
 A/Gene: msp
 C/Function:
 A/Description: inhibits gastrointestinal motility and gastric acid secretion
 C/Superfamily: spasmolytic protein; trefoil homology
 C/Keywords: duplication; hormone; pancreas
 F/1-22/Domain: signal sequence (fragment) #status predicted <SIG>
 F/23-128/Product: spasmolytic protein #status predicted <MAT>
 F/30-71/Domain: trefoil homology <TRF1>
 F/80-120/Domain: trefoil homology <TRF2>
 F/28-126,30-57,41-56,51-68,78-106,90-105,100-117/disulfide bonds: #status predicted

Query Match 25.0%; Score 5; DB 1; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5

Db 2 PRGAP 6

RESULT 25

I53637
 spasmolytic polypeptide - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C/Accession: I53637
 R/Jeffrey, G.P.; Oates, P.S.; Wang, T.C.; Babyatsky, M.W.; Brand, S.J.
 Gastroenterology 106, 336-345, 1994
 A/Title: Spasmolytic polypeptide: a trefoil peptide secreted by rat gastric mucous cells
 A/Reference number: I53637; MUID:94131228; PMID:8299900
 A/Accession: I53637
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-129 <RES>
 A/Cross-references: UNIPROT:Q90930; GB:M97255; NID:G207049; PIDN:AAA19025.1; PID:G207050
 C/Superfamily: spasmolytic protein; trefoil homology
 F/31-72/Domain: trefoil homology <TRF1>
 F/81-121/Domain: trefoil homology <TRF2>

Query Match 25.0%; Score 5; DB 1; Length 129;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5

Db 3 PRGAP 7

RESULT 26

A57325
 C-X-C chemokine LIX - mouse
 N/Alternate names: GARG-8/LIX; glucocorticocoid-attenuated response gene 8
 C/Species: Mus musculus (house mouse)
 C/Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 09-Jul-2004
 C/Accession: A57325
 R/Smith, J.B.; Herschman, H.R.
 J. Biol. Chem. 270, 16756-16765, 1995
 A/Title: Glucocorticoid-attenuated response genes encode intercellular mediators, includi
 A/Reference number: A57325; MUID:95348101; PMID:7622488
 A/Accession: A57325
 A/Status: preliminary; not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-132 <SMI>
 A/Cross-references: UNIPROT:P50228; GB:U27267; NID:G950158; PIDN:AAC52238.1; PID:G950159
 C/Superfamily: beta-thromboglobulin

Query Match 25.0%; Score 5; DB 2; Length 132;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LRCVC 12

Db 51 LRCVC 55

RESULT 27

G82796
 hypothetical protein XF0524 [imported] - Xylella fastidiosa (strain 9a5c)
 C/Species: Xylella fastidiosa
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C/Accession: G82796
 R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence
 Nature 406, 151-157, 2000
 A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A/Reference number: A82515; MUID:20365717; PMID:10910347
 A/Note: for a complete list of authors see reference number A59328 below
 A/Accession: G82796
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-142 <SIM>
 A/Cross-references: UNIPROT:Q9PFY2; GB:AB003900; GB:AF003849; NID:G9105366; PIDN:AAF83334
 A/Experimental source: strain 9a5c
 R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000

A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigre
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
 A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
 M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A/Reference number: A59328
 A/Contents: annotation
 C/Genetics:
 A/Gene: XF0524

Query Match 25.0%; Score 5; DB 2; Length 142;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5

Db 59 PRGAP 63

RESULT 28

T50331

```

hypothetical coiled-coil protein [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C:Accession: T50331
R:Layne, M.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Harris, D.
submitted to the EMBL Data Library, February 2000
A:Reference number: Z25062
A:Accession: T50331
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-147 <LYN>
A:Cross-references: UNIPROT:Q9P7P1; EMBL:AL157874; PIDN:CA875993.1; GSPDB:GN000607; SPDB:
A:Experimental source: strain 972h(-); cosmid c1718
C:Genetics:
A:Gene: SPDB:SPBC1718.03
A:Map position: 2

Query Match      25.0%; Score 5; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      14 MLETK 18
        |||||
Db       86 MLETK 90

RESULT 29
AH1645
hypothetical protein lin1705 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AH1645
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; UID:21537279; PMID:11679669
A:Accession: AH1645
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148 <GLA>
A:Cross-references: UNIPROT:Q2B45; GB:AL592022; PIDN:CAC96936.1; PID:g16414192; GSPDB:G
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin1705

Query Match      25.0%; Score 5; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      13 QMLET 17
        |||||
Db       11 QMLET 15

RESULT 30
AD1594
hypothetical protein lin1293 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD1594
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; UID:21537279; PMID:11679669
A:Accession: AD1594

```

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148 <GLA>
A:Cross-references: UNIPROT:Q92C98; GB:AL592022; PIDN:CAC96524.1; PID:g16413766; GSPDB:G
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin1293

Query Match      25.0%; Score 5; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      13 QMLET 17
        |||||
Db       11 QMLET 15

RESULT 31
D75295
DNA topology modulation protein Flak-related protein - Deinococcus radiodurans (strain R
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: D75295
R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; UID:20036896; PMID:10567266
A:Accession: D75295
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-162 <WHI>
A:Cross-references: UNIPROT:Q9RS54; GB:AE002059; GB:AE000513; NID:g6460070; PIDN:AAF1181
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2273
A:Map position: 1

Query Match      25.0%; Score 5; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      13 QMLET 17
        |||||
Db      133 QMLET 137

RESULT 32
F82743
hypothetical protein XF0947 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: F82743
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: AB2515; UID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: F82743
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-165 <SIM>
A:Cross-references: UNIPROT:Q9PFT1; GB:AE003933; GB:AE003849; NID:g9105863; PIDN:AAF8375
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir
as-Neto, E.; Docena, C.; El-Dorfy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GeneBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima J.P.; Krieger, J.E.; Kuranae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.,

```

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaak
A.; Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0947

Query Match 25.0%; Score 5; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WLRGV 11
Db 144 WLRGV 148

RESULT 33
AH1106
hypothetical protein lmc0255 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AH1106
R;Glaser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitourman, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1106
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-167 <GLA>
A;Cross-references: UNIPROT:Q8YA39; GB:NC_003210; PIDN:CAD00782.1; PID:g16409620; GSPDB:
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmc0255

Query Match 25.0%; Score 5; DB 2; Length 167;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKF 19
Db 110 LETKF 114

RESULT 34
H81294
probable lipoprotein Cj1483c [imported] - Campylobacter jejuni (strain NCTC 11168)
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: H81294
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chillin
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A;Reference number: AB1250; MUID:20150912; PMID:10686204
A;Accession: H81294
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-173 <PAR>
A;Cross-references: UNIPROT:Q9PMH9; GB:AL139078; GB:AL111168; NID:g9686723; PIDN:CAB7390
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: Cj1483c
C;Superfamily: Campylobacter jejuni probable lipoprotein Cj1483c

Query Match 25.0%; Score 5; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 MLETK 18
Db 1 MLETK 5

RESULT 35
S58124
hypothetical protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C;Accession: S58124
R;Blasczyk, R.; Loehiger, C.; Wehling, J.
submitted to the EMBL Data Library, July 1995
A;Reference number: S58124
A;Accession: S58124
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-179 <BLA>
A;Cross-references: UNIPROT:Q29839; EMBL:X89707; NID:g1052702; PIDN:CAA61854.1; PID:g1052
C;Genetics: 89/2
A;Introns: 89/2

Query Match 25.0%; Score 5; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 15 PRGAP 19

RESULT 36
D63061
precorrin-8W decarboxylase - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 16-Aug-2004
C;Accession: D63061
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; F
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.;
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi
A;Reference number: A63000; MUID:98037514; PMID:9371463
A;Accession: D63061
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-192 <MTH>
A;Cross-references: UNIPROT:O26249; GB:AE000803; GB:AE000666; NID:g2621179; PIDN:AAB84652
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH146
A;Start codon: TTG
C;Superfamily: Precorrin-6Y methyltransferase (decarboxylating) CblT

Query Match 25.0%; Score 5; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKF 19
Db 137 LETKF 141

RESULT 37
T46483
hypothetical protein DKFZp434B205.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46483
R;Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23035
A;Accession: T46483
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-192 <AAA>
A;Cross-references: UNIPROT:Q969U6; EMBL:AL137631
A;Experimental source: adult testis; clone DKFPz434B205
C;Genetics:
A;Note: DKFPz434B205.1

Query Match 25.0%; Score 5; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETK 18
|||||
DB 137 MLETK 141

RESULT 38
AH1855
hypothetical protein all0393 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AH1855
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, Nakazaki, N.; Shimpco, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH1855
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-193 <KUR>
A;Cross-references: UNIPROT:Q9Y2R4; GB:BA000019; PIDN:BA072351.1; PID: g17129738; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0393

Query Match 25.0%; Score 5; DB 2; Length 193;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETKFL 20
|||||
DB 114 ETKFL 118

RESULT 39
SAVLDM
delta large antigen - hepatitis delta virus (strain Japanese N-1)
N;Alternate names: HDag
C;Species: hepatitis delta virus
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C;Accession: A36409
R;Imazeki, F.; Omata, M.; Ohto, M.
J. Virol. 64, 5594-5599, 1990
A;Title: Heterogeneity and evolution rates of delta virus RNA sequences.
A;Reference number: A36409; MUID:91012805; PMID:2214027
A;Accession: A36409
A;Molecule type: genomic RNA
A;Residues: 1-195 <IMA>
A;Cross-references: UNIPROT:P25881; GB:D90190; GB:M58299; NID:g221691; PIDN:BAA14214.1;
C;Superfamily: hepatitis delta virus large antigen
C;Keywords: core protein
F;2-195/Product: delta large antigen #status predicted <MAT>

Query Match 25.0%; Score 5; DB 1; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5

DB 159 PRGAP 163
|||||

RESULT 40
SAVLDS
delta large antigen - hepatitis delta virus (isolate Japanese S-1)
N;Alternate names: HDag
C;Species: hepatitis delta virus
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C;Accession: B36409; SI8678
R;Imazeki, F.; Omata, M.; Ohto, M.
J. Virol. 64, 5594-5599, 1990
A;Title: Heterogeneity and evolution rates of delta virus RNA sequences.
A;Reference number: A36409; MUID:91012805; PMID:2214027
A;Accession: B36409
A;Molecule type: genomic RNA
A;Residues: 1-195 <IMA1>
A;Cross-references: UNIPROT:P25883; GB:D90192; GB:M58303; NID:g221695; PIDN:BAA14216.1;
A;Experimental source: isolate Japanese S-1
R;Imazeki, F.; Omata, M.; Ohto, M.
Nucleic Acids Res. 19, 5439, 1991
A;Title: Complete nucleotide sequence of hepatitis delta virus RNA in Japan.
A;Reference number: SI8678; MUID:92020244; PMID:1923832
A;Accession: SI8678
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: genomic RNA
A;Residues: 1-195 <IMA2>
A;Cross-references: EMBL:X60193; NID:g59497; PIDN:CAA42749.1; PID:g59498
A;Experimental source: isolate Japanese S-1
A;Note: note this sequence was submitted to the EMBL Data Library, June 1991
C;Superfamily: hepatitis delta virus large antigen
C;Keywords: core protein
F;2-195/Product: delta large antigen #status predicted <MAT>

Query Match 25.0%; Score 5; DB 1; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
|||||
DB 159 PRGAP 163

RESULT 41
QOBEC5
HOLF2 protein precursor - human cytomegalovirus (strain AD169)
N;Alternate names: hypothetical protein US2
C;Species: human cytomegalovirus, human herpesvirus 5
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004
C;Accession: E26078; S09916
R;Weston, K.; Barrell, B.G.
J. Mol. Biol. 192, 177-208, 1986
A;Title: Sequence of the short unique region, short repeats, and part of the long repeat
A;Reference number: A92935; MUID:87169717; PMID:3031311
A;Accession: E26078
A;Molecule type: DNA
A;Residues: 1-199 <WES>
A;Cross-references: UNIPROT:P09713; EMBL:X04650; NID:g59801; PIDN:CAB37096.1; PID:g44561
A;Experimental source: strain AD169
R;Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; I
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A;Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A;Reference number: S09749; MUID:90269039; PMID:2161319
A;Accession: S09916
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-199 <CHE>
A;Cross-references: EMBL:X17403; NID:g59591; PIDN:CAA35313.1; PID:g1780933
A;Experimental source: strain AD169
A;Note: this sequence was submitted to the EMBL Data Library, December 1989
C;Genetics:

A;Gene: HOLF2
 C;Superfamily: cytomegalovirus HOLF2 protein
 C;Keywords: glycoprotein; transmembrane protein
 F;1-25/Domain: signal sequence #status predicted <SIG>
 F;26-199/Product: hypothetical protein US2 #status predicted <WAT>
 F;161-186/Domain: transmembrane #status predicted <TMM>
 F;68/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 25.0%; Score 5; DB 1; Length 139;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WLRVCV 11
 |||||
 Db 130 WLRVCV 134

RESULT 42
 I46688
 complement component C8 gamma subunit - rabbit
 C;Species: Oryctolagus cuniculus (domestic rabbit)
 C;Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 09-Jul-2004
 C;Accession: I46688
 R;White, R.V.; Kaufman, K.M.; Letson, C.S.; Plattborze, P.L.; Sodetz, J.M.
 J. Immunol. 152, 2501-2508, 1994
 A;Title: Characterization of rabbit complement component C8: Functional evidence for the
 A;Reference number: I46686; MUID:94179833; PMID:7510745
 A;Accession: I46688
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-202 <WHI>
 A;Cross-references: UNIPROT:Q28679; GB:L26979; NID:g469064; PIDN:AAA31193.1; PID:g469065
 C;Superfamily: lipocalin; lipocalin; lipocalin homology
 F;42-168/Domain: lipocalin homology <LIP>

Query Match 25.0%; Score 5; DB 2; Length 202;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
 |||||
 Db 26 PRGAP 30

RESULT 43
 A97342
 nudix (MutT) family hydrolase [imported] - Clostridium acetobutylicum
 C;Species: Clostridium acetobutylicum
 C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
 C;Accession: A97342
 R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
 A;Reference number: A96900; MUID:21359325; PMID:21359325
 A;Accession: A97342
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-202 <KUR>
 A;Cross-references: UNIPROT:Q97D79; GB:AE001437; PIDN:AAK81524.1; PID:gl5026699; GSPDB:G
 A;Experimental source: Clostridium acetobutylicum ATCC824
 C;Genetics:
 A;Gene: CAC3601

Query Match 25.0%; Score 5; DB 2; Length 202;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETKFL 20
 |||||
 Db 14 ETKFL 18

RESULT 44
 D70935
 hypothetical protein Rv2170 - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C;Accession: D70935
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: D70935
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-206 <COL>
 A;Cross-references: UNIPROT:O53504; GB:AL021957; GB:AL123456; NID:g3242293; PIDN:CAA17474
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: Rv2170

Query Match 25.0%; Score 5; DB 2; Length 206;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMWL 8
 |||||
 Db 39 APMWL 43

RESULT 45
 S28710
 RNA-directed RNA polymerase homolog - sugar beet yellows virus (fragment)
 C;Species: sugar beet yellows virus, SBV
 C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 08-Oct-1999
 C;Accession: S28710
 R;Agranovsky, A.A.; Boyko, V.P.; Karasev, A.V.; Lunina, N.A.; Koonin, E.V.; Dolja, V.V.
 J. Gen. Virol. 72, 15-23, 1991
 A;Title: Nucleotide sequence of the 3'-terminal half of beet yellows closterovirus RNA ge
 A;Reference number: S28710; MUID:91116305; PMID:1990061
 A;Accession: S28710
 A;Molecule type: DNA
 A;Residues: 1-218 <AGR>
 A;Cross-references: EMBL:X53462; NID:g58878; PIDN:CRA37549.1; PID:g808947

Query Match 25.0%; Score 5; DB 2; Length 218;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETKFL 20
 |||||
 Db 64 ETKFL 68

RESULT 46
 S33204
 hypothetical protein - long-stalked stitchwort
 C;Species: Stellaria longipes (long-stalked stitchwort)
 C;Date: 09-Jun-1994 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
 C;Accession: S33204
 R;Zhang, X.H.
 submitted to the EMBL Data Library, April 1993
 A;Reference number: S33204
 A;Accession: S33204
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-222 <ZHA>
 A;Cross-references: UNIPROT:Q41348; EMBL:X71601; NID:g297175; PID:g297176
 C;Superfamily: hypothetical protein H11647

Query Match 25.0%; Score 5; DB 2; Length 222;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;

```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRGAP 5
    |||||
Db 54 PRGAP 58

RESULT 47
B95910
conserved hypothetical protein SMB20806 [imported] - Sinorhizobium meliloti (strain 1021)
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: B95910
R:Finan, T.M.; Weidner, K.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481331
A:Accession: B95910
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-229 <KUR>
A:Cross-references: UNIPROT:Q92W01; GB:AL591985; PIDN:CAC48946.1; PID:GL5140431; GSPDB:G
A:Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.M.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
A:Genetics:
C:Gene: SMB20806
A:Genome: plasmid

Query Match 25.0%; Score 5; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 ETKFL 20
    |||||
Db 115 ETKFL 119

RESULT 48
B81429
hypothetical protein Cj0122 [imported] - Campylobacter jejuni (strain NCTC 11168)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: B81429
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: B81429
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-229 <PAR>
A:Cross-references: UNIPROT:Q9PJ05; GB:ALJ39074; GB:ALJ11168; NID:G6967505; PIDN:CAB7260
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj0122

Query Match 25.0%; Score 5; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LETKF 19
    |||||
Db 186 LETKF 190

```

RESULT 49

B90222

conserved hypothetical protein [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C:Accession: B90222

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozeza, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.

arett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A39139

A:Accession: B90222

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-233 <KUR>

A:Cross-references: UNIPROT:Q9UVC5; GB:AE006641; NID:GI3813905; PIDN:AAK41033.1; GSPDB:G

C:Genetics:

A:Gene: SSO0737

Query Match

Best Local Similarity 100.0%; Pred. No. 2.2e+02; Length 233;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

14 MLETK 18

|||||

Db

95 MLETK 99

RESULT 50

T29697

hypothetical protein F31A3.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T29697

R:Murray, J.; Lee, T.T.

submitted to the EMBL Data Library, May 1996

A:Description: The sequence of C. elegans cosmid F31A3.

A:Reference number: Z20667

A:Accession: T29697

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-239 <MUR>

A:Cross-references: UNIPROT:Q19917; EMBL:U58742; PIDN:AAB36858.1; GSPDB:CN00028; CBSP:

A:Experimental source: strain Bristol N2; clone F31A3

C:Genetics:

A:Gene: CBSP:F31A3.4

A:Map position: X

A:Introns: 57/2; 132/2

Query Match

Best Local Similarity 100.0%; Pred. No. 2.3e+02; Length 239;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

15 LETKF 19

|||||

Db

74 LETKF 78

RESULT 51

T20657

hypothetical protein F09B8.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T20657

R:Percy, C.

submitted to the EMBL Data Library, May 1996

A:Reference number: Z19307

A:Accession: T20657

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-240 <WIL>

A;Cross-references: UNIPROT:Q19268; EMBL:Z73896; PIDN:CAA98056.1; GSPDB:GN00022; CBSP:FD
 A;Experimental source: clone F09B8
 C;Genetics:
 A;Gene: CBSP:F09B8.1
 A;Map position: 4
 A;Introns: 18/1; 54/1; 97/1; 149/1

Query Match 25.0%; Score 5; DB 2; Length 240;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKF 19
 |||||
 DB 89 LETKF 93
 |||||

RESULT 52
 B99833
 Hypothetical protein SA0589 [imported] - Staphylococcus aureus (strain N315)
 C;Species: Staphylococcus aureus
 C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 16-Aug-2004
 C;Accession: B99833
 R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogino,
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A;Reference number: A89758; MUID:21311952; PMID:11418146
 A;Accession: B99833
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-247 <KUR>
 A;Cross-references: UNIPROT:Q59VV2; GB:BA000018; PID:g13700524; PIDN:BAB41821.1; GSPDB:G
 A;Experimental source: strain N315
 C;Genetics:
 A;Gene: SA0589
 C;Superfamily: ATP-binding cassette homology

Query Match 25.0%; Score 5; DB 2; Length 247;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETK 18
 |||||
 DB 1 MLETK 5
 |||||

RESULT 53
 LNDGPS
 pulmonary surfactant protein A precursor - dog
 N;Alternate names: pulmonary surfactant 32K apoprotein; pulmonary surfactant-associated
 C;Species: Canis lupus familiaris (dog)
 C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
 C;Accession: A25296; A61227; A60142
 R;Benson, B.; Hawgood, S.; Schilling, J.; Clements, J.; Damm, D.; Cordell, B.; White, R.
 Proc. Natl. Acad. Sci. U.S.A. 82, 6379-6383, 1985
 A;Title: Structure of canine pulmonary surfactant apoprotein: cDNA and complete amino ac
 A;Reference number: A25296; MUID:86016705; PMID:3863100
 A;Accession: A25296
 A;Molecule type: mRNA
 A;Residues: 1-248 <BEN>
 A;Note: the authors translated the codon TTG for residue 60 as Pro
 A;Note: part of the sequence, including the amino end of the mature protein, was confirm
 R;Liau, D.F.; Ryan, S.F.
 Chem. Phys. Lipids 59, 29-38, 1991
 A;Title: Purification of surfactant protein A from dog lung by reconstitution with surfa
 A;Reference number: A61227; MUID:92163993; PMID:1790579
 A;Accession: A61227
 A;Molecule type: protein
 A;Residues: 18-32 <LIA>
 R;Ross, G.F.; Meuth, J.; Ohning, B.; Kim, Y.; Whitsett, J.A.
 Biochim. Biophys. Acta 870, 267-278, 1986
 A;Title: Purification of canine surfactant-associated glycoproteins A. Identification of

A;Reference number: A60142; MUID:86159848; PMID:3006781
 A;Accession: A60142
 A;Molecule type: protein
 A;Residues: 24-34;95-101,'X',103-108 <ROS>
 R;Patthy, L.
 Nature 325, 490, 1987
 A;Reference number: A93388; MUID:87115834; PMID:3808053
 A;Contents: annotation; animal lectin domain homology
 C;Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers th
 C;Comment: This protein is a sialoglycoprotein synthesized by alveolar type II cells. It
 pendent on the presence of calcium ions.
 C;Superfamily: mannose-binding lectin; C-type lectin homology
 C;Keywords: alveolar proteinosis; calcium; disulfide bond; gaseous exchange; hydroxyproli
 F;1-17/Domain: signal sequence #status predicted <SIG>
 F;18-248/Product: pulmonary surfactant protein A #status experimental <MPT>
 F;28-102/Region: collagen-like
 F;127-246/Domain: C-type lectin homology <LCH>
 F;20,207/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;30/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match 25.0%; Score 5; DB 1; Length 248;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 MWLRC 10
 |||||
 DB 1 MWLRC 5
 |||||

RESULT 54
 T05424
 hypothetical protein F28A23.120 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 C;Accession: T05424
 R;Bevan, M.; Weichselgartner, M.; Fartmann, B.; Granderath, K.; Dauner, D.; Herzl, A.; Ne
 submitted to the Protein Sequence Database, October 1998
 A;Reference number: Z15415
 A;Accession: T05424
 A;Molecule type: DNA
 A;Residues: 1-249 <BEV>
 A;Cross-references: UNIPROT:Q49493; EMBL:AL021961
 A;Experimental source: cultivar Columbia; BAC clone F28A23
 C;Genetics:
 A;Map position: 4
 A;Introns: 21/3; 101/1; 122/3; 139/1; 154/2; 219/2
 A;Note: F28A23.120

Query Match 25.0%; Score 5; DB 2; Length 249;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKF 19
 |||||
 DB 222 LETKF 226
 |||||

RESULT 55
 C70823
 Hypothetical protein RV0736 - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C;Accession: C70823
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, C.; Churcher, C.; Harris, D.; Gordon, S.;
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A;Authors: Scares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: C70823
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA

A;Residues: 1-266 <STO>
A;Cross-references: UNIPROT:Q9KA34; GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BAB061
A;Experimental source: strain C-125
C;Genetics:
A;Gene: fliH

Query Match 25.0%; Score 5; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKF 19
|||||

Db 233 LETKF 237

RESULT 61
T00114
exodeoxyribonuclease III (EC 3.1.1.1.2) - Actinobacillus actinomycetemcomitans
N;Alternate names: deoxynuclease
C;Species: Actinobacillus actinomycetemcomitans
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004
C;Accession: T00114
R;Nakano, Y.; Yoshida, Y.; Yamashita, Y.; Koga, T.
Biochim. Biophys. Acta 1442, 409-414, 1998
A;Title: A gene cluster for 6-deoxy-L-talan synthesis in Actinobacillus actinomycetemcomitans
A;Reference number: Z14111; MUID:99023768; PMID:9805002
A;Accession: T00114
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-267 <NAK>
A;Cross-references: UNIPROT:O66261; EMBL:AB010415; NID:g3132248; PIDN:BAA28143.1; PID:g3132248
A;Experimental source: strain NCTC 9710
C;Function:
A;Description: catalyzes degradation of double-stranded DNA; removes damaged DNA at cytoskeleton
C;Superfamily: exodeoxyribonuclease III
C;Keywords: DNA repair; hydrolase

Query Match 25.0%; Score 5; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKF 19
|||||

Db 96 LETKF 100

RESULT 62
T21429
hypothetical protein F26H11.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21429
R;Barlow, K.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19421
A;Accession: T21429
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-273 <WIL>
A;Cross-references: EMBL:Z81515; PIDN:CAB04193.1; GSPDB:GN00020; CESP:F26H11.5
A;Experimental source: clone F26H11
C;Genetics:
A;Gene: CESP:F26H11.5
A;Map position: 2
A;Introns: 5/3; 159/3

Query Match 25.0%; Score 5; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETKFL 20
|||||

Db 176 ETKFL 180

RESULT 63
S16563
nolG protein - Rhizobium meliloti
C;Species: Rhizobium meliloti
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 08-Oct-1999
C;Accession: S16563
R;Baev, N.; Endre, G.; Petrovics, G.; Banfalvi, Z.; Kondorosi, A.
Mol. Gen. Genet. 228, 113-124, 1991
A;Title: Six nodulation genes of nod box locus 4 in Rhizobium meliloti are involved in nif gene expression
A;Reference number: S16561; MUID:91360053; PMID:1909418
A;Accession: S16563
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-277 <BAE>
A;Cross-references: EMBL:X58632; NID:g46331; PIDN:CAA41487.1; PID:g46334

Query Match 25.0%; Score 5; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKF 19
|||||

Db 115 LETKF 119

RESULT 64
C69854
conserved hypothetical protein Y3QC - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: C69854
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, A.; Ehrlich, S.D.; Emmer, P.T.; Entian, K.D.; Errington, J.; Fabbet, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallercia, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoisappel, S.; Hosono, S.; Hullo, M.F.; Koster, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lotters, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelie, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, I.; Winters, P.; Wipat, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, A.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: C69854
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-278 <KUN>
A;Cross-references: UNIPROT:O34423; GB:Z99110; GB:AL009126; NID:g2633472; PIDN:CAB13106.1
A;Experimental source: strain 168
C;Genetics:
A;Gene: Y3QC

Query Match 25.0%; Score 5; DB 2; Length 278;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMWL 8
|||||

Db 264 APMWL 268

RESULT 65
T27610
hypothetical protein ZC477.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T20906
R;McMurray, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19345
A;Accession: T20906
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-305 <WIL>
A;Cross-references: UNIPROT:O17805; EMBL:Z81503; PIDN:CAB04111.1; GSPDB:GN00021; CESP:F147
A;Experimental source: clone F147
C;Genetics:
A;Gene: CESP:F14F7.1
A;Map position: 3
A;Introns: 27/3; 49/3

Query Match 25.0%; Score 5; DB 2; Length 305;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 193 PRGAP 197

RESULT 71
AD2202
transcription regulator all3171 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AD2202
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2202
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-306 <KUR>
A;Cross-references: UNIPROT:Q9YS88; GB:BA000019; PIDN:BA074870.1; PID:gl17132266; GSPDB:G000000000
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all3171

Query Match 25.0%; Score 5; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ETKFL 20
Db 107 ETKFL 111

RESULT 72
H75049
hypothetical protein PAB1459 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: H75049
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
A;Reference number: A75001
A;Accession: H75049
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-308 <KAW>
A;Cross-references: UNIPROT:Q9UYW9; GB:AJ248287; GB:AL096836; NID:g5458657; PIDN:CAB5029
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB1459

C;Superfamily: Pyrococcus abyssi hypothetical protein PAB1459

Query Match 25.0%; Score 5; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
Db 181 RGAPM 185

RESULT 73
G71110
hypothetical protein PH0654 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C;Accession: G71110
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon Pyrococcus horikoshii
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: G71110
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-308 <KAW>
A;Cross-references: UNIPROT:O58388; GB:AP000003; NID:g3236130; PIDN:BAA29745.1; PID:g3257
A;Experimental source: strain OT3
A;Note: This accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH0654

Query Match 25.0%; Score 5; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPM 6
Db 181 RGAPM 185

RESULT 74
T49591
probable 35 kDa ribonuclease H [imported] - Neurospora crassa
N;Alternate names: protein B3E4.40
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49591
R;Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-317 <SCH>
A;Cross-references: UNIPROT:Q9P5X8; EMBL:AL355931; GSPDB:GN00116; NCSP:B3E4.40
A;Experimental source: BAC clone B3E4; strain OR74A
C;Genetics:
A;Gene: NCSP:B3E4.40
A;Map position: 6
A;Introns: 223/1

Query Match 25.0%; Score 5; DB 2; Length 317;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETK 18
Db 264 MLETK 268

RESULT 75
T33845

hypothetical protein Y44E3A.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33845
R:Woessner, J.; Graves, T.; Keppler, D.
submitted to the EMBL Data Library, November 1998
A:Description: The sequence of C. elegans cosmid Y44E3A.
A:Reference number: Z21422
A:Accession: T33845
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-318 <WOE>
A:Cross-references: UNIPROT:Q9TXZ0; EMBL:AF106589; PIDN:AAC78231.1; GSPDB:GN00019; CESP:
A:Experimental source: strain Bristol N2; clone Y44E3A
C:Genetics:
A:Gene: CESP:Y44E3A.5
A:Map position: 1
A:Introns: 6/3; 41/1; 155/3; 223/3

Query Match 25.0%; Score 5; DB 2; Length 318;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LETKF 19
|
|
|
|
Db 167 LETKF 171

Search completed: October 26, 2004, 07:23:39
Job time : 19.5 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:48:06 ; Search time 91.5 Seconds
(without alignments)
125.765 Million cell updates/sec

Title: US-10-066-965A-4

Perfect score: 20

Sequence: 1 PRGAPWMLRCVCQMLETKFL 20

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1825181 seqs, 575374646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Uniprot_02.*

1: uniprot_sprot.*

2: uniprot_treml.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	35.0	82	Q8VJN1	Q8VJN1 mycobacteri
2	6	30.0	179	TXN9_BOVIN	Q18883 bos taurus
3	6	30.0	188	Q8TB70	Q8TB70 homo sapien
4	6	30.0	211	Q87727	Q87727 vibrio algi
5	6	30.0	224	Q9EZC2	Q9EZC2 staphylococ
6	6	30.0	226	TXN9_HUMAN	Q14530 homo sapien
7	6	30.0	226	TXN9_MOUSE	Q9CQ79 mus musculu
8	6	30.0	226	TXN9_RAT	Q8K581 rattus norv
9	6	30.0	226	CAG33216	CAG33216 homo sapi
10	6	30.0	238	P73223	P73223 synechocyst
11	6	30.0	250	Q7P4Z8	Q7P4Z8 fusbacteri
12	6	30.0	253	Q6HK94	Q6HK94 bacillus th
13	6	30.0	253	Q7JA41	Q7JA41 bacillus ce
14	6	30.0	253	Q8IS20	Q8IS20 bacillus an
15	6	30.0	253	AA340868	AA340868 bacillus
16	6	30.0	253	AAT30974	AAT30974 bacillus
17	6	30.0	259	Q6P9X7	Q6P9X7 rattus norv
18	6	30.0	259	AAH60541	AAH60541 rattus no
19	6	30.0	267	AAH70183	AAH70183 homo sapi
20	6	30.0	271	AROE_CLOPE	Q8XMI8 clostridium
21	6	30.0	278	AROE_BACHD	Q8K933 bacillus ha
22	6	30.0	315	Q9VLD6	Q9VLD6 drosophila
23	6	30.0	319	Q810R3	Q810R3 drosophila
24	6	30.0	384	Q9C280	Q9C280 neurospora
25	6	30.0	388	Q6F387	Q6F387 homo sapien
26	6	30.0	388	AAH64143	AAH64143 homo sapi
27	6	30.0	395	Q96MA7	Q96MA7 homo sapien
28	6	30.0	421	RSMB_XLFA	Q9PEV0 xylella fas
29	6	30.0	426	RSMB_VIBVU	Q8DD55 vibrio vuln
30	6	30.0	426	RSMB_VIBVU	Q7MGK4 vibrio vuln
31	6	30.0	427	RSMB_VIBPA	Q87KD3 vibrio para

32	6	30.0	431	1	RSMB_XYLFT	Q87AR1 xylella fas
33	6	30.0	431	2	Q7S5V2	Q7S5V2 neurospora
34	6	30.0	431	2	Q6LLJ1	Q6LLJ1 photobacter
35	6	30.0	431	2	CAG21837	CAG21837 photobact
36	6	30.0	434	1	RSMB_VIBCH	Q9KVU5 vibrio chol
37	6	30.0	436	1	RSMB_XANAC	Q9G222 xanthomonas
38	6	30.0	437	1	RSMB_XANCP	Q8P4G1 xanthomonas
39	6	30.0	454	2	Q6NE86	Q6NE86 neurospora
40	6	30.0	454	2	CAF06276	CAF06276 neurospor
41	6	30.0	483	2	Q8HZR5	Q8HZR5 potorous tr
42	6	30.0	490	2	Q95U34	Q95U34 drosophila
43	6	30.0	583	1	AP1_KLULA	P56095 kluyveromyc
44	6	30.0	593	2	Q8R7Q4	Q8R7Q4 thermoanaer
45	6	30.0	672	2	Q6FK03	Q6FK03 candida gla
46	6	30.0	772	2	Q7P394	Q7P394 fusbacteri
47	6	30.0	772	2	Q8RGC0	Q8RGC0 fusbacteri
48	6	30.0	793	2	Q93325	Q93325 caenorhabdi
49	6	30.0	887	2	Q8STF6	Q8STF6 caenorhabdi
50	6	30.0	1037	2	Q7YV13	Q7YV13 trypanosoma
51	6	30.0	1345	2	Q8A180	Q8A180 bacteroides
52	6	30.0	1502	2	Q9UPZ6	Q9UPZ6 homo sapien
53	6	30.0	1524	2	Q7RCM2	Q7RCM2 plasmodium
54	6	30.0	2516	2	Q84BD6	Q84BD6 myxococcus
55	6	30.0	3396	2	O62263	O62263 caenorhabdi
56	5	25.0	23	2	Q83SX5	Q83SX5 salmonella
57	5	25.0	28	2	Q8KE5	Q8KE5 rabies viru
58	5	25.0	28	2	Q8KE7	Q8KE7 rabies viru
59	5	25.0	28	2	Q8KE9	Q8KE9 rabies viru
60	5	25.0	28	2	Q8KF0	Q8KF0 rabies viru
61	5	25.0	28	2	Q8KF2	Q8KF2 rabies viru
62	5	25.0	28	2	Q8KF3	Q8KF3 rabies viru
63	5	25.0	28	2	Q8KF4	Q8KF4 rabies viru
64	5	25.0	28	2	Q8KF5	Q8KF5 rabies viru
65	5	25.0	28	2	Q8KF6	Q8KF6 rabies viru
66	5	25.0	28	2	Q8KF8	Q8KF8 rabies viru
67	5	25.0	28	2	Q8KF9	Q8KF9 rabies viru
68	5	25.0	28	2	Q8KG1	Q8KG1 rabies viru
69	5	25.0	28	2	Q8KG2	Q8KG2 rabies viru
70	5	25.0	28	2	Q8KG3	Q8KG3 rabies viru
71	5	25.0	28	2	Q8KG4	Q8KG4 rabies viru
72	5	25.0	28	2	Q8KG7	Q8KG7 rabies viru
73	5	25.0	28	2	Q8KG8	Q8KG8 rabies viru
74	5	25.0	28	2	Q8KG9	Q8KG9 rabies viru
75	5	25.0	28	2	Q8KH0	Q8KH0 rabies viru
76	5	25.0	28	2	Q8KH1	Q8KH1 rabies viru
77	5	25.0	28	2	Q8KH9	Q8KH9 rabies viru
78	5	25.0	28	2	Q8KK1	Q8KK1 rabies viru
79	5	25.0	28	2	Q8KK2	Q8KK2 rabies viru
80	5	25.0	28	2	Q8KK4	Q8KK4 rabies viru
81	5	25.0	28	2	Q8KK5	Q8KK5 rabies viru
82	5	25.0	28	2	Q8KK6	Q8KK6 rabies viru
83	5	25.0	28	2	Q8KK7	Q8KK7 rabies viru
84	5	25.0	28	2	Q8KK8	Q8KK8 rabies viru
85	5	25.0	28	2	Q8KK9	Q8KK9 rabies viru
86	5	25.0	28	2	Q8KL0	Q8KL0 rabies viru
87	5	25.0	28	2	Q8KL1	Q8KL1 rabies viru
88	5	25.0	28	2	Q8KL2	Q8KL2 rabies viru
89	5	25.0	28	2	Q8KL3	Q8KL3 rabies viru
90	5	25.0	28	2	Q8KL5	Q8KL5 rabies viru
91	5	25.0	28	2	Q9IWC4	Q9IWC4 rabies viru
92	5	25.0	28	2	Q9IWC5	Q9IWC5 rabies viru
93	5	25.0	28	2	Q9QD10	Q9QD10 rabies viru
94	5	25.0	28	2	Q9QD11	Q9QD11 rabies viru
95	5	25.0	28	2	Q9QD12	Q9QD12 rabies viru
96	5	25.0	28	2	Q9QD13	Q9QD13 rabies viru
97	5	25.0	28	2	Q9QD14	Q9QD14 rabies viru
98	5	25.0	28	2	Q9QD15	Q9QD15 rabies viru
99	5	25.0	28	2	Q9QD16	Q9QD16 rabies viru
100	5	25.0	28	2	Q9QD17	Q9QD17 rabies viru

ALIGNMENTS

```
RESULT 1
Q8VJN1 ID Q8VJN1 PRELIMINARY; PRT; 82 AA.
AC Q8VJN1
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein M2341.
GN OrderedLocustNames=M2341;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolorav J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W.R., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
DR EMBL; AB007077; AAK4625.1; -.
DR TIGR; MT2341; -.
KW Hypothetical protein.
SQ SEQUENCE 82 AA; 8995 MW; C46009E4E66DC09E CRC64;

Query Match 35.0%; Score 7; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGAPMWL 8
Db 6 RGAPMWL 12

RESULT 2
TXN9 BOVIN ID TXN9 BOVIN STANDARD; PRT; 179 AA.
AC O18683;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Thioredoxin domain containing protein 9 (Protein 1-4) (Fragment).
GN Name=TXNDC9; Synonyms=APACD;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Corpus luteum;
RA Brule S., Lussier J.G.;
RT "Bovine ATP binding protein.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the UPF0071 family. Some similarity to the
CC thioredoxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF027733; AAB84006.1; -.
```

```
DR InterPro; IPR006662; ThioRed.
DR InterPro; IPR006663; ThioRedox_dom2.
DR Pfam; PF00085; ThioRedoxin; 1.
FT NON_TER 1
SQ SEQUENCE 179 AA; 21051 MW; 687318949AD494C CRC64;

Query Match 30.0%; Score 6; DB 1; Length 179;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
Db 49 LETKFL 54

RESULT 3
Q8TB70 ID Q8TB70 PRELIMINARY; PRT; 188 AA.
AC Q8TB70;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE APACD protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024223; AAB42423.2; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006662; ThioRed.
DR InterPro; IPR006663; ThioRedox_dom2.
DR Pfam; PF00085; ThioRedoxin; 1.
KW Redox-active center.
SQ SEQUENCE 188 AA; 22199 MW; B375C22E69DF6523 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
Db 119 LETKFL 124
```


RESULT 4

O87727
 ID O87727 PRELIMINARY; PRT; 211 AA.
 AC O87727;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Fmu protein.
 GN Name=Fmu;
 OS Vibrio alginolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=663;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=138-2;
 RX MEDLINE=98386520; PubMed=9720051;
 RA Nakamura T., Yamamoto N., Stumpe S., Unemoto T., Bakker E.P.;
 RT "Cloning of the trkAH gene cluster and characterization of the Trk K+-
 RT uptake system of Vibrio alginolyticus.";
 RL Microbiology 144:2281-2289(1998).
 DR EMBL; D86411; BAA1226.1; -;
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001678; Fmu_NOL1/Nop2p.
 DR InterPro; IPR006027; NusB.
 DR InterPro; IPR006174; NusB_region.
 DR Pfam; PF01189; Noll_Nop2_Fmu; 1.
 DR Pfam; PF01029; NusB; 1.
 DR ProDom; PD005242; NusB_region; 1.
 SQ SEQUENCE 211 AA; 24135 MW; 2B39B568BA78B5C6 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 211;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 6; Conservative 0; Mismatches 0; Gaps 0; Indels 0;
 QY 4 APMWLR 9
 Db 169 APMWLR 174

RESULT 5

O9EZC2
 ID O9EZC2 PRELIMINARY; PRT; 224 AA.
 AC O9EZC2;
 DT 01-WAR-2001 (TrEMBLrel. 16, Created)
 DT 01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-WAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE ISI182 transposase.
 OS Staphylococcus intermedius.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1285;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MLS-2;
 RX MEDLINE=21154722; PubMed=11230937;
 RA Boerlin P., Burnens A.P., Frey J., Kuhnert P., Nicolet J.;
 RT "Molecular epidemiology and genetic linkage of macrolide and
 RT aminoglycoside resistance in Staphylococcus intermedius of canine
 RT origin."
 RL Vet. Microbiol. 79:155-169(2001).
 DR EMBL; AF292922; AAG42229.1; -;
 SQ SEQUENCE 224 AA; 26309 MW; C0E3400DDC9B8F1D CRC64;

Query Match 30.0%; Score 6; DB 2; Length 224;
 Best Local Similarity 100.0%; Pred. No. 96;
 Matches 6; Conservative 0; Mismatches 0; Gaps 0; Indels 0;
 QY 13 QMLETK 18
 Db 174 QMLETK 179

RESULT 6

TXN9 HUMAN
 ID TXN9 HUMAN STANDARD; PRT; 226 AA.
 AC O14530; Q6NSFS; Q9BRU6;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Thioredoxin domain containing protein 9 (protein 1-4) (ATP binding
 DE protein associated with cell differentiation).
 GN Name=TXNDC9; Synonyms=APACD;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leukemia;
 RA Shiosaka T.;
 RT "Differential expression of 1-4 gene in functionally distinct ME-1
 RT subclones.";
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
 RT "Cloning of human full open reading frames in Gateway(TM) system entry
 RT vector (pDONR201).";
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow, Brain, and Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Stapleton M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Not known, associated with cell differentiation.
 CC -!- SIMILARITY: Belongs to the UPF0071 family. Some similarity to the
 CC thioredoxin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB006679; BAA21881.1; -;
 DR EMBL; CR456935; CAG33216.1; -;
 DR EMBL; BC005968; AAH05968.1; -;
 DR EMBL; BC022864; AAH22864.1; -;
 DR EMBL; BC070183; AAH70183.1; ALT_INIT.
 DR Genew; HGNC:24110; TXNDC9.
 DR InterPro; IPR006662; Thiored.
 DR InterPro; IPR006663; Thioredox_dom2.
 DR Pfam; PF00085; Thioredoxin; 1.

```

FT CONFLICT 122 122 K -> N (in Ref. 1).
FT CONFLICT 214 214 I -> M (in Ref. 1).
SQ SEQUENCE 226 AA; 26534 MW; 2735A2562D1143C3 CRC64;

Query Match 30.0%; Score 6; DB 1; Length 226;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
Db 119 LETKFL 124

RESULT 7
TXN9_MOUSE STANDARD; PRT; 226 AA.
AC Q9CQ79;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DE Thiredoxin domain containing protein 9 (ATP binding protein
DE associated with cell differentiation).
GN Name=txndc9; Synonyms=Apac9;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grinstead S., Guscinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Ienhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sadelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Datchenko K., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Baha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

```

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.B., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Smerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Not known, associated with cell differentiation (By
CC similarity).
CC -!- SIMILARITY: Belongs to the UPF0071 family. Some similarity to the
CC thioredoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AK002893; BAB22438.1; -
CC EMBL; AK006170; BAB24440.1; -
CC EMBL; AK010709; BAB27134.1; -
CC EMBL; AK011424; BAB27611.1; -
CC EMBL; AK016756; BAB30412.1; -
CC EMBL; AK028525; BAC25991.1; -
CC EMBL; BC022947; AAB22947.1; -
CC MGD; MGI:2138153; Apac9.
CC InterPro; IPR006662; Thiredox.
CC InterPro; IPR006663; Thiredox_dom2.
CC Pfam; PF00085; Thiredoxin; 1.
SQ SEQUENCE 226 AA; 26260 MW; B90255DDAE402BC CRC64;

Query Match 30.0%; Score 6; DB 1; Length 226;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
Db 119 LETKFL 124

RESULT 8
TXN9_RAT STANDARD; PRT; 226 AA.
AC Q8K581;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Thiredoxin domain containing protein 9 (ES cell-related protein).
GN Name=txndc9;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RT Li Y., Xu C., Zhang Y.;
RT "Cloning and analysis of up-regulated genes in rat liver."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Not known, associated with cell differentiation (By
CC similarity).
CC -!- SIMILARITY: Belongs to the UPF0071 family. Some similarity to the
CC thioredoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF508022; AM34684.1; ALT_INIT.
DR InterPro; IPR006663; Rhoredox_Gom2.
SQ SEQUENCE 226 AA; 26272 MW; EDASBF54C19AE69A CRC64;

Query Match 30.0%; Score 6; DB 1; Length 226;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
Db 119 LETKFL 124

RESULT 9
CAG33216 PRELIMINARY; PRT; 226 AA.

ID CAG33216;
AC CAG33216;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE APACD protein.
GN APACD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RT "Cloning of human full open reading frames in Gateway(TM) system entry
RT vector (pDONR201)";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR456935; CAG33216.1; -;
SQ SEQUENCE 226 AA; 26534 MW; 2735A2562D1143C3 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
Db 119 LETKFL 124

RESULT 10
P73223 PRELIMINARY; PRT; 238 AA.

ID P73223;
AC P73223;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE S11925 protein.
GN OrderedLocusNames=s111925;
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;

RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RX MEDLINE=97061201; PubMed=8905231;

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Hatao A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90904; BAA17250.1; -;
DR FIR; S75336; S75336.
KW Complete proteome.
SQ SEQUENCE 238 AA; 25843 MW; F016A42E3DD3E588 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
Db 165 LETKFL 170

RESULT 11
Q7P4Z8 PRELIMINARY; PRT; 250 AA.

ID Q7P4Z8;
AC Q7P4Z8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Shikimate 5-dehydrogenase (EC 1.1.1.25).
GN Name=PNV0648;
OS Fusobacterium nucleatum subsp. vincentii ATCC 49256.
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=209882;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 49256;
RA Karpatal V., Ivanova N., Anderson I., Reznik G., Bhattacharyya A.,
RA Gardner W.L., Mikhailova N., Larsen N., D'Souza M., Walunas T.,
RA Haselkorn R., Overbeek R., Kyrpides N.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

DR EMBL; AABF0100090; EAA23737.1; -;
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004764; F:shikimate 5-dehydrogenase activity; IEA.
DR GO; GO:0016089; P:aromatic amino acid family biosynthesis, sh...; IEA.
DR InterPro; IPR006151; Shikimate_DH.
DR Pfam; PF01488; Shikimate_DH; 1.
KW Oxidoreductase.

SQ SEQUENCE 250 AA; 28872 MW; 7FFE3D1EE8750DDA CRC64;
Query Match 30.0%; Score 6; DB 2; Length 250;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
Db 187 LETKFL 192

RESULT 12
Q6HK94 PRELIMINARY; PRT; 253 AA.

ID Q6HK94;
AC Q6HK94;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.

GN ORFNames=BT9727.1700;
 OS Bacillus thuringiensis serovar konkukian str. 97-27.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OC Bacillus thuringiensis serovar Konkukian.
 OX NCBI_TaxID=281309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=97-27;
 RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
 RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
 RA Richardson P., Rubin E., Tice H.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017355; AAT59607.1; --
 DR InterPro: IPR004183; LiGB.
 DR Pfam: PF02900; LiGB; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 253 AA; 28619 MW; 863533F0058A872E CRC64;

Query Match 30.0%; Score 6; DB 2; Length 253;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETKF 19
 |||||
 Db 87 MLETKF 92

RESULT 13
 Q73A41 PRELIMINARY; PRT; 253 AA.
 AC Q73A41;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Oxidoreductase.
 GN OrderedLocusNames=BCE1944;
 OS Bacillus cereus (strain ATCC 10987).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=222523;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14960714;
 RA Rasko D.A., Ravel J., Okstad O.A., Helgason E., Cer R.Z., Jiang L.,
 RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
 RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
 RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic adaptations and a large plasmid related to Bacillus anthracis pXO1.";
 RL Nucleic Acids Res. 32:977-988(2004).
 DR EMBL; AE017270; AAS40868.1; --
 DR TIGR; BCE1944; --
 DR InterPro: IPR004183; LiGB.
 DR Pfam: PF02900; LiGB; 1.
 KW Complete proteome.
 SQ SEQUENCE 253 AA; 28543 MW; C38AC1D894E451BA CRC64;

Query Match 30.0%; Score 6; DB 2; Length 253;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETKF 19
 |||||
 Db 87 MLETKF 92

RESULT 14
 Q81S20 PRELIMINARY; PRT; 253 AA.
 AC Q81S20; Q610A0; Q6KU74;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Oxidoreductase.
 GN OrderedLocusNames=BA1859, BAS1723; ORFNames=GBAA1859;

OS Bacillus anthracis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1392;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ames / isolate Porton;
 RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
 RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
 RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
 RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
 RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
 RA DeBoy R.T., Madpu R., Baugherty S.C., Durkin A.S., Haft D.H.,
 RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
 RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
 RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
 RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
 RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
 RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
 RT "The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria.";
 RL Nature 423:81-86(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ames / isolate 0581;
 RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
 RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
 RA Fraser C.M.;
 RT "Bacillus anthracis comparative genomics.";
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Stearne;
 RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
 RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
 RA Richardson P., Rubin E., Tice H.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017030; AAP25762.1; --
 DR EMBL; AE017334; AAT30974.1; --
 DR EMBL; AE017225; AAT54039.1; --
 DR TIGR; BA1859; --
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro: IPR004183; LiGB.
 DR Pfam: PF02900; LiGB; 1.
 SQ SEQUENCE 253 AA; 28572 MW; 3EC117D7768D3076 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 253;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETKF 19
 |||||
 Db 87 MLETKF 92

RESULT 15
 AAS40868 PRELIMINARY; PRT; 253 AA.
 ID AAS40868
 AC AAS40868;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 04-MAY-2004 (TrEMBLrel. 27, Last annotation update)
 DE Oxidoreductase.
 GN BCE1944.
 OS Bacillus cereus (strain ATCC 10987).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=222523;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14960714;
 RA Rasko D.A., Ravel J., Okstad O.A., Helgason E., Cer R.Z., Jiang L.,
 RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,

```

RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
RT adaptations and a large plasmid related to Bacillus anthracis pX01.";
RL Nucleic Acids Res. 32:977-988(2004).
DR EMBL: AB017270; AAS40868.1; -.
DR TIGR: BCE1944; -.
SQ SEQUENCE 253 AA; 28543 MW; C38AC1D894E451BA CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 253;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETKF 19
Db 87 MLETKF 92

RESULT 16
AAT30974
ID AAT30974 PRELIMINARY; PRT; 253 AA.
AC AAT30974;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Oxidoreductase.
GN GBAA1859.
OS Bacillus anthracis str. Ames 0581.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group; Bacillus anthracis.
OX NCBI_TaxID=261594;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S., Fraser C.M.;
RT "Bacillus anthracis comparative genomics.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB017334; AAT30974.1; -.
SQ SEQUENCE 253 AA; 28572 MW; 3EC117D7768D3076 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 253;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETKF 19
Db 87 MLETKF 92

RESULT 17
Q6P9X7
ID Q6P9X7 PRELIMINARY; PRT; 259 AA.
AC Q6P9X7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE LOC280671 protein (Fragment).
GN Names=LOC280671;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary gland;
RA Strausberg R.;

```

```

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary gland;
RA Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC060541; AAH60541.1; -.
DR InterPro: IPR006663; Thioresox_dom2.
FT NON TER 1
SQ SEQUENCE 259 AA; 29764 MW; 008F5E943F78D68A CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
Db 152 LETKFL 157

RESULT 18
AAH60541
ID AAH60541 PRELIMINARY; PRT; 259 AA.
AC AAH60541;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE LOC280671 protein (Fragment).
GN LOC280671.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary gland;
RA Strausberg R.;

```



```

RL Nucleic Acids Res. 28:4317-4331(2000).
CC -!- CATALYTIC ACTIVITY: Shikimate + NADP(+) = 5-dehydroshikimate +
CC NADPH.
CC
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC fourth step.
CC
CC -!- SIMILARITY: Belongs to the shikimate dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: AP001511; BAB05043.1; -
CC FSP; D83815; D83815.
CC HSP; Q58484; INVT.
CC HAMAP; MF_00222; -; 1.
CC InterPro; IPR006151; Shikimate DH.
CC Pfam; PF01488; Shikimate DH; 1.
CC TIGRFAMS; TIGR00507; aroE_1.
CC TIGRFAMS; TIGR01809; Shik-DH-AROM; 1.
CC Aromatic amino acid biosynthesis; Complete proteome; NADP;
KW Oxidoreductase.
SQ SEQUENCE 278 AA; 30630 MW; 2000144C84686264 CRC64;

Query Match 30.0%; Score 6; DB 1; Length 278;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
Db 225 LETKFL 230
|||||

RESULT 22
Q9VLD6 PRELIMINARY; PRT; 315 AA.
AC Q9VLD6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE CG4454-PA (IDJ36125p).
GN ORFNames=CG4454;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champé M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abell J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

```

```

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murphy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.X., Wasman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett W., Wotley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RT Science 287:2185-2195(2000).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.K., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RN SEQUENCE FROM N.A.
RG FLYBASE;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RN SEQUENCE FROM N.A.
RG FLYBASE;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [7]
RN SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003624; AAF52757.1; -.
DR EMBL; AY061432; AAL28980.1; -.
DR FlyBase; FBgn0032105; CG4454.

```


SQ SEQUENCE 315 AA; 34680 MW; ACIDF9F5960C9F2A CRC64;

Query Match 30.0%; Score 6; DB 2; Length 315;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LETKFL 20
 |||||
 Db 40 LETKFL 45

RESULT 23

Q810R3 PRELIMINARY; PRT; 319 AA.

AC Q810R3;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 23, Last annotation update)
 DE CG4454-PB (RE31802p).
 GN ORFNames=CG4454;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazee R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew K.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Chervy J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foslter K., Gabriellista A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
 RA Patrick J.N., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,

RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
 RT melanogaster euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskaas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RG FLYBASE;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RG FLYBASE;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacle J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003624; AAN10690.1; --
 DR EMBL; BT001609; AAN71364.1; --
 DR FlyBase; FBgn0032105; CG4454.
 SQ SEQUENCE 319 AA; 35115 MW; 1F70AA48CDC2D369 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 319;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LETKFL 20
 |||||
 Db 40 LETKFL 45

RESULT 24

Q9C280 PRELIMINARY; PRT; 384 AA.

AC Q9C280;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Hypothetical protein B13A5.030.
 GN Name=B13A5.030;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]

RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto K., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Mutsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara T., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Hisigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45(2004).
DR EMBL; AK057264; BAB71401.1; -. 2B7BEB535377F87 CRC64;
SQ SEQUENCE 395 AA; 43443 MW; 2B7BEB535377F87 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 395;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLETK 18
|||||
Db 365 QMLETK 370

RESULT 28
ID RSBM XYLF6 STANDARD; PRT; 421 AA.
AC Q9P9V0;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DE Ribosomal RNA small subunit methyltransferase B (EC 2.1.1.-) (rRNA
DE (cytosine-C(5)-)-methyltransferase) (16S rRNA m5C967
DE methyltransferase).
GN Name=rsmB; Synonyms=rmb; OrderedLocusNames=Xf0928;
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5C;
RX MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Colombo A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Couteau N.B., Camargo C., Costa F.P., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.O.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,

RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Moon C.H., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moun D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.D.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa";
RL Nature 406:151-159(2000).
CC -!- FUNCTION: Specifically methylates the cytosine at position 967
CC (m5C967) of 16S rRNA (By similarity).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC homocysteine + rRNA containing C(5)-methylcytosine.
CC -!- SIMILARITY: Belongs to the methyltransferase superfamily. RsmB/NOP
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB003932; AAF83738.1; -.
DR PIR; A82744; A82744.
DR InterPro; IPR004573; Fm1 mtfase.
DR InterPro; IPR001678; Fm1 NOLL/Nop2p.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR006027; NusB.
DR InterPro; IPR006174; NusB_region.
DR InterPro; IPR000051; SAM_Bind.
DR Pfam; PF01189; Noll_Nop2_Sun; 1.
DR Pfam; PF01029; NusBj_1.
DR ProDom; PD005242; NusB_dom; 1.
DR TIGRPFAMs; TIGR00563; rsmB; 1.
DR PROSITE; PS01153; NOLL_NOP2_SUN; 1.
KW Complete proteome; Methyltransferase; rRNA processing; Transferase.
SQ SEQUENCE 421 AA; 45909 MW; 2F0BB6B697E6F387 CRC64;

Query Match 30.0%; Score 6; DB 1; Length 421;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMWLR 9
|||||
Db 159 APMWLR 164

RESULT 29
ID RSBM VIBVU STANDARD; PRT; 426 AA.
AC Q8DD5;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DE Ribosomal RNA small subunit methyltransferase B (EC 2.1.1.-) (rRNA
DE (cytosine-C(5)-)-methyltransferase) (16S rRNA m5C967
DE methyltransferase).
GN Name=rsmB; Synonyms=rmb; OrderedLocusNames=VV11046;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;

```

RA Rhee J.H., Kim S.-Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.B.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Specifically methylates the cytosine at position 967
CC (m5C967) of 16S rRNA (By similarity).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC homocysteine + rRNA containing C(5)-methylcytosine.
CC -!- SIMILARITY: Belongs to the methyltransferase superfamily. RsmB/NOP
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AR016800; AAC09533.1; -.
CC InterPro; IPR004573; Fmu_mtfase.
CC InterPro; IPR001678; Fmu_NOL1/Nop2p.
CC InterPro; IPR006027; NusB.
CC InterPro; IPR006174; NusB region.
CC InterPro; IPR000051; SAM_bind.
CC Pfam; PF01189; Noll_Nop2_Sun; 1.
CC Pfam; PF01029; NusB; 1.
CC ProDom; PD005242; NusB_dom; 1.
CC TIGRFAMs; TIGR00563; rsmB; 1.
CC PROSITE; PS01153; NOLL_NOP2_SUN; 1.
CC Complete proteome; Methyltransferase; rRNA processing; Transferase.
CC SEQUENCE 426 AA; 47677 MW; D53200B63B05CDFC CRC64;
SQ

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP005343; BAC95991.1; -.
CC InterPro; IPR004573; Fmu_mtfase.
CC InterPro; IPR001678; Fmu_NOL1/Nop2p.
CC InterPro; IPR006027; NusB.
CC InterPro; IPR006174; NusB region.
CC InterPro; IPR000051; SAM_bind.
CC Pfam; PF01189; Noll_Nop2_Sun; 1.
CC Pfam; PF01029; NusB; 1.
CC ProDom; PD005242; NusB_dom; 1.
CC TIGRFAMs; TIGR00563; rsmB; 1.
CC PROSITE; PS01153; NOLL_NOP2_SUN; 1.
CC Complete proteome; Methyltransferase; rRNA processing; Transferase.
CC SEQUENCE 426 AA; 47687 MW; D5320602762A1DFC CRC64;
SQ

```

```

Query Match          30.0%; Score 6; DB 1; Length 426;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4  APWMLR 9
        |||||
Db      169  APWMLR 174

```

```

RESULT 31
RSMB_VIBPA
ID   RSMB_VIBPA          STANDARD;          PRT;   427 AA.
AC   Q87KD3;
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DE   Ribosomal RNA small subunit methyltransferase B (EC 2.1.1.-) (rRNA
DE   (cytosine-C(5)-)-methyltransferase) (16S rRNA m5C967
DE   methyltransferase).
DE   Name=rsmB; Synonyms=rsmB; OrderedLocusNames=VF3044;
OS   Vibrio parahaemolyticus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=670;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RMD 2210633 / Serotype O3:K6;
RX   MEDLINE=22508454; PubMed=12620739;
RA   Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA   Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA   Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT   "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT   distinct from that of V. cholerae."
RL   Lancet 361:743-749(2003).
CC   -!- FUNCTION: Specifically methylates the cytosine at position 967
CC (m5C967) of 16S rRNA (By similarity).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC homocysteine + rRNA containing C(5)-methylcytosine.
CC -!- SIMILARITY: Belongs to the methyltransferase superfamily. RsmB/NOP
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP005083; BAC61307.1; -.
CC InterPro; IPR004573; Fmu_mtfase.

```

```

DR InterPro; IPR001678; Fnu_NOL1/Nop2p.
DR InterPro; IPR006027; NusB.
DR InterPro; IPR000051; SAM bind.
DR Pfam; PF01189; Noll_Nop2_Sun; 1.
DR Pfam; PF01029; NusB; 1.
DR PROSITE; PS01153; NOLL_NOP2_SUN; 1.
DR PROSITE; PS01029; NusB; 1.
DR PROSITE; PS01153; NOLL_NOP2_SUN; 1.
DR Complete proteome; Methyltransferase; rRNA processing; Transferase.
KW SEQUENCE 427 AA; 47835 MW; 653C0C335D1D6B5C CRC64;

Query Match 30.0%; Score 6; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMWLR 9
Db 169 APMWLR 174

RESULT 32
RSMB_XYLFT STANDARD; PRT; 431 AA.
AC Q87AR1;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ribosomal RNA small subunit methyltransferase B (EC 2.1.1.-) (rRNA
DE (cytosine-C(5)-)-methyltransferase) (16S rRNA m5C967
DE methyltransferase).
GN Name=rsmB; Synonyms=rsmB; OrderedLocusNames=PD1761;
OS Xylella fastidiosa (strain Temeculal / AFCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22421331; PubMed=12533478;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Tsai S.M.,
RA Carrer H., Carriaro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kurama E.B.,
RA Marino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA Baia G.S., Bianco S.R., Brito M.S., Camavan F.S., Celestino A.V.,
RA da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sassaki F.T., Sena J.A.D.,
RA de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa."
RL J. Bacteriol. 185:1018-1026(2003).
CC -!- FUNCTION: Specifically methylates the cytosine at position 967
CC (m5C967) of 16S rRNA (by similarity).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC homocysteine + rRNA containing C(5)-methylcytosine.
CC -!- SIMILARITY: Belongs to the methyltransferase superfamily. RsmB/NOP
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE012559; AAO29595.1; -.
CC InterPro; IPR004573; Fnu mtfrase.
DR InterPro; IPR001678; Fnu_NOL1/Nop2p.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR006027; NusB.

```

```

DR InterPro; IPR000051; SAM bind.
DR Pfam; PF01189; Noll_Nop2_Sun; 1.
DR Pfam; PF01029; NusB; 1.
DR PROSITE; PS01153; NOLL_NOP2_SUN; 1.
DR PROSITE; PS01029; NusB; 1.
DR Complete proteome; Methyltransferase; rRNA processing; Transferase.
KW SEQUENCE 431 AA; 46932 MW; 35C7C763B8B0908B CRC64;

Query Match 30.0%; Score 6; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMWLR 9
Db 167 APMWLR 172

RESULT 33
Q7SSV2 PRELIMINARY; PRT; 431 AA.
AC Q7SSV2;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU05614.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehm B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley J., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kanal M., Kamysovellis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000339; EAA30904.1; -.
SQ SEQUENCE 431 AA; 48408 MW; 9E466BBFFB0D173B CRC64;

Query Match 30.0%; Score 6; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
Db 399 LETKFL 404

RESULT 34
Q6LLJ1 PRELIMINARY; PRT; 431 AA.
AC Q6LLJ1;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Putative Sun protein.
GN Name=T4096; OrderedLocusNames=PBPA3581;
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).

```

```
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP SEQUENCE FROM N.A.
RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
RA Cestaro A., Malacrida G., Simionati B., Cannata N., Bartlett D.,
RA Valle G.;
RT "Genome analysis of Photobacterium profundum reveals the complexity of
RT high pressure adaptations.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR378674; CAG21837.1; -.
DR InterPro; IPR004573; Fmu_mtfase.
DR InterPro; IPR001678; Fmu_NOL1/Nop2p.
DR InterPro; IPR006027; NusB.
DR InterPro; IPR006174; NusB_region.
DR InterPro; IPR001638; SBP_Dac_3.
DR Pfam; PF01189; Noll_Nop2_Fmu; 1.
DR Pfam; PF01029; NusB; 1.
DR ProDom; PD005242; NusB_region; 1.
DR TIGRFAMs; TIGR00563; rsmB; 1.
DR PROSITE; PS01153; NOLL_NOP2_SUN; 1.
DR PROSITE; PS01039; SBP_BACTERIAL_3; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 431 AA; 48054 MW; 1B5D728A0D557688 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMWLR 9
Db 169 APMWLR 174
```

```
RESULT 35
CAG21837
ID CAG21837 PRELIMINARY; PRT; 431 AA.
AC CAG21837;
DT 10-MAY-2004 (TrEMBLrel. 27, Created)
DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative Sun protein.
GN T4096 OR PBPR3581.
OS Photobacterium profundum (Photobacterium sp. (strain S99)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP SEQUENCE FROM N.A.
RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
RA Cestaro A., Malacrida G., Simionati B., Cannata N., Bartlett D.,
RA Valle G.;
RT "Genome Analysis of Photobacterium profundum reveals the complexity of
RT high pressure adaptations.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ss9;
RA Cestaro A.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR378674; CAG21837.1; -.
SQ SEQUENCE 431 AA; 48054 MW; 1B5D728A0D557688 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMWLR 9
Db 169 APMWLR 174
```

```
RESULT 36
RSMB_VIBCH
ID RSMB_VIBCH STANDARD; PRT; 434 AA.
AC Q9KV05;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Ribosomal RNA small subunit methyltransferase B (EC 2.1.1.1.-) (rRNA
DE (cytosine-C(5)-)-methyltransferase) (16S rRNA m5C967
DE methyltransferase).
DE Name=rsmB; Synonyms=rsmB; OrderedLocusNames=VC00044;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483 (2000).
CC -!- FUNCTION: Specifically methylates the cytosine at position 967
CC (m5C967) of 16S rRNA (By similarity).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC homocysteine + rRNA containing C(5)-methylcytosine.
CC -!- SIMILARITY: Belongs to the methyltransferase superfamily. RsmB/NOP
CC subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AF004095; AAF93222.1; -.
CC FIR; G82372; G82372.
CC TIGR; VC0044; -.
CC InterPro; IPR004573; Fmu_mtfase.
CC InterPro; IPR001678; Fmu_NOL1/Nop2p.
CC InterPro; IPR006027; NusB.
CC InterPro; IPR006174; NusB_region.
CC InterPro; IPR000051; SAM_Bind.
CC Pfam; PF01189; Noll_Nop2_Sun; 1.
CC Pfam; PF01029; NusB; 1.
CC ProDom; PD005242; NusB_dom; 1.
CC TIGRFAMs; TIGR00563; rsmB; 1.
CC PROSITE; PS01153; NOLL_NOP2_SUN; 1.
KW Complete proteome; Methyltransferase; rRNA processing; Transferase.
SQ SEQUENCE 434 AA; 48340 MW; D03DE106F9F85ADD CRC64;

Query Match 30.0%; Score 6; DB 1; Length 434;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMWLR 9
Db 177 APMWLR 182
```

```
RESULT 37
RSMB_XANAC
ID RSMB_XANAC STANDARD; PRT; 436 AA.
```

Q8PG22;
 05-JUL-2004 (Rel. 44, Last sequence update)
 05-JUL-2004 (Rel. 44, Last annotation update)
 05-JUL-2004 (Rel. 44, Last annotation update)
 Ribosomal RNA small subunit methyltransferase B (EC 2.1.1.1.-) (rRNA
 (cytosine-C(5)-)-methyltransferase) (16S rRNA m5C967
 methyltransferase).
 Name=rmbB; Synonyms=rmbB, sun; OrderedLocusNames=XAC3799;
 Xanthomonas axonopodis (pv. citri).
 Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 Xanthomonadaceae; Xanthomonas.
 NCBI_TaxID=92829;
 [1]
 SEQUENCE FROM N.A.
 STRAIN=306 / ATCC 13902 / XV 101;
 MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
 da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 Canarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
 Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 Spínola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Terza R.I.D.,
 Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 Setubal J.C., Kitajima J.P.;
 "Comparison of the genomes of two Xanthomonas pathogens with differing
 host specificities";
 Nature 417:459-463 (2002).
 CC -!- FUNCTION: Specifically methylates the cytosine at position 967
 (m5C967) of 16S rRNA (by similarity).
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
 homocysteine + rRNA containing C(5)-methylcytosine.
 CC -!- SIMILARITY: Belongs to the methyltransferase superfamily. RmbB/NOP
 subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE012028; AAM38641.1; -.
 CC InterPro; IPR004573; Fnu mtfrase.
 CC InterPro; IPR001678; Fnu NOL1/Nop2p.
 CC InterPro; IPR006027; NusB.
 CC InterPro; IPR006174; NusB region.
 CC Pfam; PF01189; Noll_Nop2_Sun; 1.
 CC Pfam; PF01029; NusB; 1.
 CC ProDom; PD005242; NusB_dom; 1.
 CC TIGRFAMs; TIGR00563; rmbB; 1.
 CC PROSITE; PS01153; NOLL_NOP2_SUN; 1.
 CC Complete proteome; Methyltransferase; rRNA processing; Transferase.
 KW SEQUENCE 436 AA; 47034 MW; 02FF031ED1D04068 CRC64;

Query Match 30.0%; Score 6; DB 1; Length 436;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMWLR 9

Db 172 APMWLR 177

RESULT 38

RSMB_XANCP
 ID RSMB_XANCP STANDARD; PRT; 437 AA.
 AC Q8P4G1;
 DT 05-JUL-2004 (Rel. 44, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Ribosomal RNA small subunit methyltransferase B (EC 2.1.1.1.-) (rRNA
 (cytosine-C(5)-)-methyltransferase) (16S rRNA m5C967
 methyltransferase).
 DE Name=rmbB; Synonyms=rmbB, sun; OrderedLocusNames=XCC3747;
 Xanthomonas campestris (pv. campestris).
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
 da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 Canarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
 Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 Spínola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Terza R.I.D.,
 Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 Setubal J.C., Kitajima J.P.;
 "Comparison of the genomes of two Xanthomonas pathogens with differing
 host specificities";
 Nature 417:459-463 (2002).
 CC -!- FUNCTION: Specifically methylates the cytosine at position 967
 (m5C967) of 16S rRNA (by similarity).
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
 homocysteine + rRNA containing C(5)-methylcytosine.
 CC -!- SIMILARITY: Belongs to the methyltransferase superfamily. RmbB/NOP
 subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE012495; AAM43004.1; -.
 CC InterPro; IPR004573; Fnu mtfrase.
 CC InterPro; IPR001678; Fnu NOL1/Nop2p.
 CC InterPro; IPR006027; NusB.
 CC InterPro; IPR006174; NusB region.
 CC Pfam; PF01189; Noll_Nop2_Sun; 1.
 CC Pfam; PF01029; NusB; 1.
 CC ProDom; PD005242; NusB_dom; 1.
 CC TIGRFAMs; TIGR00563; rmbB; 1.
 CC PROSITE; PS01153; NOLL_NOP2_SUN; 1.
 CC Complete proteome; Methyltransferase; rRNA processing; Transferase.
 KW SEQUENCE 437 AA; 47204 MW; 15F2FF80700440B3 CRC64;

Query Match 30.0%; Score 6; DB 1; Length 437;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 APMWLR 9

Db 173 APMWLR 178

```

RESULT 39
Q6NE86 PRELIMINARY; PRT; 454 AA.
ID Q6NE86;
AC Q6NE86;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein 45B12.090.
GN Name=45B12.090;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX295538; CAF06276.1; -.
KW Hypothetical protein.
SQ SEQUENCE 454 AA; 50847 MW; 52DF1FA686F940D8 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 454;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
    |||||
DB 422 LETKFL 427

RESULT 40
CAF06276 PRELIMINARY; PRT; 454 AA.
ID CAF06276;
AC CAF06276;
DT 02-MAR-2004 (TRENBLrel. 27, Created)
DT 02-MAR-2004 (TRENBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein 45B12.090.
GN 45B12.090
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX295538; CAF06276.1; -.
KW Hypothetical protein.
SQ SEQUENCE 454 AA; 50847 MW; 52DF1FA686F940D8 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 454;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
    |||||
DB 422 LETKFL 427

RESULT 41
Q8H2R5 PRELIMINARY; PRT; 483 AA.
ID Q8H2R5;
AC Q8H2R5;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Keratin 7.
OS Potorous tridactylus (Potoroo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Potoroidae; Potorous.
OX NCBI_TaxID=9310;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22246069; PubMed=12359226;
RA Smith F.J.D., Porter R.M., Corden L.D., Lunny D.P., Lane E.B.,
RA McLean W.H.I.;
RT "Cloning of human, murine, and marsupial keratin 7 and a survey of K7
RT expression in the mouse.";
RL Biochem. Biophys. Res. Commun. 297:818-827(2002).
CC -!- SIMILARITY: Belongs to the intermediate filament family.
DR EMBL; AF509889; AAN64033.1; -.
DR HSP; P08670; 1GK4.
DR GO; GO:0005882; C:intermediate filament; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR011000; AOPolp_III_like.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin I.
DR InterPro; IPR003054; Keratin II.
DR Pfam; PF00038; Filament; 1.
DR PRINTS; PRO1248; TYPE1KERATIN.
DR PRINTS; PRO1276; TYPE2KERATIN.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Keratin.
SQ SEQUENCE 483 AA; 53286 MW; E3632B3267517BCE CRC64;

Query Match 30.0%; Score 6; DB 2; Length 483;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLETK 18
    |||||
DB 134 QMLETK 139

RESULT 42
Q95U34 PRELIMINARY; PRT; 490 AA.
ID Q95U34; Q9VSS5;
AC Q95U34; Q9VSS5;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE GH11113p (CG5288-PA) (CG5288-pb).
GN ORFNames=CG5288;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunco J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

```


CC or send an email to license@sib-sib.ch).

```

CC EMBL; AF006499; AAC39320.1; -.
DR HSP; F03069; 2DGC.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00217; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; 1.
KW Activator; DNA-binding; Nuclear protein; Transcription regulation.
FT DNA BIND 38 67 Basic motif.
FT DOMAIN 79 107 Leucine-zipper.
SQ SEQUENCE 583 AA; 63885 MW; A4A96B12D7F5C983 CRC64;

Query Match 30.0%; Score 6; DB 1; Length 583;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
   |||||
Db 95 LETKFL 100

RESULT 44
Q8R7Q4 PRELIMINARY; PRT; 593 AA.
AC Q8R7Q4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted ATPase.
GN Name=LYCS2; OrderedLocusNames=TTE2344;
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4.
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of the T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013176; AM25485.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR010559; His_kin_int.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF06580; His_kinase; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_C; 1.
DR PROSITE; PS00885; HAMP; 1.
KW Complete proteome.
SQ SEQUENCE 593 AA; 69032 MW; 558213122A093559 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 593;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GAPMWL 8
   |||||
Db 155 GAPMWL 160

```

RESULT 45

```

Q6FK03 PRELIMINARY; PRT; 672 AA.
AC Q6FK03;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to sp|P28743 Saccharomyces cerevisiae YPL155c KIP2 kinesin-
DE related protein.
GN ORFNames=CAGL0M021899;
OS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG GENOLEVURES;
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrist A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Boiotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CK380959; CAG62417.1; -.
DR InterPro; IPR00194; ATPase_a/bcentre.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN 1.
DR PROSITE; PS00667; KINESIN_MOTOR_DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 672 AA; 73932 MW; 17832DA3A3EA98FE CRC64;

Query Match 30.0%; Score 6; DB 2; Length 672;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MLETKF 19
   |||||
Db 612 MLETKF 617

RESULT 46
Q7P394 PRELIMINARY; PRT; 772 AA.
AC Q7P394;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical membrane spanning protein.
GN Name=ENV2117;
OS Fusobacterium nucleatum subsp. vincentii ATCC 49256.
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=209882;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 49256;
RA Karpatrial V., Ivanova N., Anderson I., Reznik G., Bhattacharyya A.,
RA Gardner W.L., Mikhailova N., Larsen N., D'Souza M., Walunas T.,
RA Haselkorn R., Overbeek R., Kyrpides N.;

```

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AABP01000003; EAA25145.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 772 AA; 90058 MW; 4B861956C1F51058 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 772;
 Best Local Similarity 100.0%; Pred.No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
 |||||
 Db 300 LETKFL 305

RESULT 47
 Q8RGCO PRELIMINARY; PRT; 772 AA.
 ID Q8RGCO;
 AC Q8RGCO;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein FNO384.
 GN OrderedLocustNames=FNO384;
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
 OC Fusobacterium.
 OX NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25586;
 RX MEDLINE=21886394; PubMed=11889109;
 RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasileva O., Chu L., Kogan X., Chaga O., Goltzman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fontstein M., Kyriades N.C., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 RT nucleatum strain ATCC 25586.";
 RL J. Bacteriol. 184:2005-2018(2002).
 DR EMBL; AE010549; AAL94587.1; --
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 772 AA; 90123 MW; 9B3A35DF7FFCB213 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 772;
 Best Local Similarity 100.0%; Pred.No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
 |||||
 Db 300 LETKFL 305

RESULT 48
 Q93325 PRELIMINARY; PRT; 793 AA.
 ID Q93325; Q9XZQ6;
 AC Q93325; Q9XZQ6;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein C33G3.1a (DYC-1 protein) (Fragment).
 GN Name=C33G3.1a; Synonyms=dyc-1;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;

RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Matthews L.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N2;
 RX MEDLINE=20453785; PubMed=10996789;
 RA Gieseler K., Segalat L., Grisoni K.;
 RT "Genetic suppression of phenotypes arising from mutations in
 RT dystrophin-related genes in Caenorhabditis elegans.";
 RL Curr. Biol. 10:1092-1097(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N2;
 RA segalat L.S.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z78540; CAB01738.2; --
 DR EMBL; AJ133838; CAB44432.1; --
 DR PIR; TI9690; TI9690.
 DR WormPep; C33G3.1a; CE24825.
 DR InterPro; IPR011036; PH related.
 DR Pfam; PF00640; PID; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 793 AA; 88624 MW; C83701F4873E14B1 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 793;
 Best Local Similarity 100.0%; Pred.No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLETK 18
 |||||
 Db 403 QMLETK 408

RESULT 49
 Q8STF6 PRELIMINARY; PRT; 887 AA.
 ID Q8STF6;
 AC Q8STF6;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein C33G3.1b.
 GN Name=C33G3.1b;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RL "Genome sequence of the nematode C.elegans: A platform for
 RL investigating biology.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Gajadaty S.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Matthews L.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z72501; CAD30430.1; --
 DR EMBL; Z78540; CAD30430.1; JOINED.

DR EMBL; Z72501; CAD30431.1; JOINED.
 DR EMBL; Z78540; CAD30431.1; -.
 DR IntAct; Q8STF6; -. CR30500.
 DR WormPep; C33G3.1b; PH related.
 DR InterPro; IPR011036; PH related.
 DR InterPro; IPR006020; PTB_PID.
 DR Pfam; PF00640; PID; 1.
 DR SMART; SM00462; PTB; 1.
 DR PROSITE; PS01179; PID; 1.
 DR Hypothetical protein.
 SW SEQUENCE 887 AA; 99567 MW; 4479D811513A9CA5 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 887;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLETK 18

DB 497 QMLETK 502

RESULT 50

Q7YV13 PRELIMINARY; PRT; 1037 AA.
 AC Q7YV13;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Protein kinase, putative.
 GN ORFNames=TB927.2.4200;
 OS Trypanosoma brucei.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22789168; PubMed=12907728;
 RA El-Sayed N.M.A., Ghedin E., Song J., MacLeod A., Bringaud F.,
 RA Larkin C., Wanless D., Peterson J., Hou L., Taylor S., Tweedie A.,
 RA Biteau N., Khalak H.G., Lin X., Mason T., Hannick L., Caler E.,
 RA Blandin G., Bartholomew D., Simpson A.J., Kaul S., Zhao H., Pai G.,
 RA Van Aken S., Uterback T., Haas B., Koo H.L., Umayam L., Suh B.,
 RA Gerrard C., Leech V., Qi R., Zhou S., Schwartz D., Feldblyum T.,
 RA Salzberg S., Tait A., Turner M.R., Ullu E., White O., Melville S.,
 RA Adams M.D., Fraser C.M., Donelson J.E.;
 RT "The sequence and analysis of Trypanosoma brucei chromosome II.";
 RL Nucleic Acids Res. 31:4856-4863(2003).
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AB017169; AAQ15877.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
 DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
 SW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 KQ SEQUENCE 1037 AA; 113326 MW; A87B8C4442745274 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 1037;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VCOMLE 16

DB 85 VCOMLE 90

RESULT 51

Q8A180

ID Q8A180 PRELIMINARY; PRT; 1345 AA.
 AC Q8A180;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Two-component system sensor histidine kinase/response regulator,
 DE hybrid ('one-component system').
 GN OrderedLocusNames=BT3786;
 OS Bacteroides thetaiotaomicron.
 OC Bacteriia; Bacteroidetes; Bacteroides (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 OX NCBI_TaxID=818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=VPI-5482 / ATCC 29148;
 RX MEDLINE=22550858; PubMed=12663928;
 RA Xu J., Bjursell M.K., Hamrod J., Deng S., Carmichael L.K.,
 RA Chiang H.C., Hooper L.V., Gordon J.I.;
 RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
 RL Science 299:2074-2076(2003).
 CC -!- SIMILARITY: Contains 1 HTH arnC/xylS-type DNA-binding domain.
 CC -!- SIMILARITY: Contains 1 histidine kinase domain.
 DR EMBL; AS016942; AAO78891.1; -.
 DR HSSP; P08402; 1B00.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0000156; F:two-component response regulator activity; IEA.
 DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
 DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
 DR InterPro; IPR003594; ATPbind_ATPase.
 DR InterPro; IPR011006; Chey_like.
 DR InterPro; IPR005467; His_kinase.
 DR InterPro; IPR003661; His_kina_N.
 DR InterPro; IPR000005; HTHArac.
 DR InterPro; IPR011110; Reg_prop.
 DR InterPro; IPR001789; Response_reg.
 DR InterPro; IPR011123; Y_Y_Y.
 DR Pfam; PF02518; HATPase_c; 1.
 DR Pfam; PF00512; Hiska; 1.
 DR Pfam; PF00165; HTH_Arac; 2.
 DR Pfam; PF07494; Reg_prop; 11.
 DR Pfam; PF00072; Response_reg; 1.
 DR Pfam; PF07495; Y_Y_Y; 1.
 DR PRINTS; PRO00032; HTHARAC.
 DR ProDom; PD000039; Response_reg; 1.
 DR SMART; SM00387; HATPase_c; 1.
 DR SMART; SM00388; Hiska; 1.
 DR SMART; SM00342; HTH_Arac; 1.
 DR SMART; SM00448; REC; 1.
 DR PROSITE; PS0109; HIS_KIN; 1.
 DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
 DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
 KW Complete proteome; DNA-binding; Kinase; Phosphorylation;
 KW Sensory transduction; Transcription regulation; Transferase.
 SQ SEQUENCE 1345 AA; 155269 MW; 7247BDA9C7FADAB CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 1345;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 VCOMLE 16
 DB 1202 VCOMLE 1207
 RESULT 52
 Q9UPZ6 PRELIMINARY; PRT; 1502 AA.
 ID Q9UPZ6
 AC Q9UPZ6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DE 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 GN KIAA0960 protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]_
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99246063; PubMed=10231032;
 RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hiroseawa M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:63-70(1999).
 DR EMBL; AB023177; BAA76804.2; -.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; TSP_1; 10.
 DR SMART; SM00209; TSP1; 15.
 DR PROSITE; PS50092; TSP1; 11.
 FT NON_TER 1
 SQ SEQUENCE 1502 AA; 168205 MW; F36BFA138FA9BFI CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 1502;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 13 QMLETK 18
 Db 1307 QMLETK 1312
 RESULT 53
 Q7RCM2 PRELIMINARY; PRT; 1524 AA.
 ID Q7RCM2;
 AC Q7RCM2;
 DT 01-MAR-2004 (TReMBLrel. 26, Created)
 DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN Name=PY05757;
 OS Plasmodium yoelii yoelii.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=73239;
 RN [1]_
 RP SEQUENCE FROM N.A.
 RC STRAIN=17XNL;
 RX PubMed=12368865;
 RA Carlton J.M., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
 RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
 RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
 RA Shallom S.J., van Aken S.E., Riedmiller S.B., Feldblyum T.V.,
 RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
 RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
 RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
 RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
 RA Carucci D.J.;
 RA "Genome sequence and comparative analysis of the model rodent malaria
 RT parasite Plasmodium yoelii yoelii.";
 RL Nature 419:512-519(2002).
 CC !- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; ABL01001868; EAA17831.1; -.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 1524 AA; 178163 MW; E2E6D84E9EFAFAD CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 1524;

Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 13 QMLETK 18
 Db 211 QMLETK 216
 RESULT 54
 Q84BD6 PRELIMINARY; PRT; 2516 AA.
 ID Q84BD6;
 AC Q84BD6;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Adventurous gliding motility protein K (Fragment).
 GN Name=agmK;
 OS Myxococcus xanthus.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
 OC Cytophasterineae; Myxococcaceae; Myxococcus.
 OX NCBI_TaxID=34;
 RN [1]_
 RP SEQUENCE FROM N.A.
 RA Hartzell P.L., Youderian P.A.;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY204471; AAC22907.1; -.
 DR InterPro; IPR001440; TPR.
 DR InterPro; IPR008941; TPR-like.
 DR Pfam; PF00515; TPR; 16.
 DR SMART; SM00028; TPR; 11.
 DR PROSITE; PS50005; TPR; 3.
 DR PROSITE; PS50293; TPR_REGION; 2.
 KW Repeat; TPR repeat.
 FT NON_TER 1
 SQ SEQUENCE 2516 AA; 280635 MW; C3D908A0E2F6CA49 CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 2516;
 Best Local Similarity 100.0%; Pred. No. 8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 13 QMLETK 18
 Db 794 QMLETK 799
 RESULT 55
 O62263 PRELIMINARY; PRT; 3396 AA.
 ID O62263;
 AC O62263;
 DT 01-AUG-1998 (TReMBLrel. 07, Created)
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Hypothetical protein F54B3.1.
 GN Name=F54B3.1;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]_
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RA "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.";
 RT Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA McMurray A.A.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z48583; CAA88472.1; -.
 DR PIR; T22613; T22613.

```
DR WormPep; F54B3.1; CE17865.
KW Hypothetical protein.
SQ SEQUENCE 3396 AA; 391879 MW; 337F20A3A9BD2CE7 CRC64;

Query Match      30.0%; Score 6; DB 2; Length 3396;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      13 QMLETK 18
Db      1962 QMLETK 1967

RESULT 56
Q83SX5 ID Q83SX5 PRELIMINARY; PRT; 23 AA.
AC Q83SX5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN Ordered locus names=t2947;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AE016844; AAO70500.1; -.
KW Hypothetical protein.
SQ SEQUENCE 23 AA; 2797 MW; 9EBB0BCB88CB36AA CRC64;

Query Match      25.0%; Score 5; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      11 VCOML 15
Db      16 VCOML 20

RESULT 57
Q80KE5 ID Q80KE5 PRELIMINARY; PRT; 28 AA.
AC Q80KE5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192397; AAO45872.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AAC3C6A CRC64;

Query Match      25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
Db      21 PRGAP 25

RESULT 58
Q80KE7 ID Q80KE7 PRELIMINARY; PRT; 28 AA.
AC Q80KE7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192395; AAO45870.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AAC3C6A CRC64;

Query Match      25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
Db      21 PRGAP 25

RESULT 59
Q80KE9 ID Q80KE9 PRELIMINARY; PRT; 28 AA.
AC Q80KE9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192393; AAO45868.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3024 MW; 2E24A7A9AAC3C6A CRC64;

Query Match      25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAP 5
Db      21 PRGAP 25

RESULT 60
Q80XFO
```

```

ID Q80KF0 PRELIMINARY; PRT; 28 AA.
AC Q80KF0;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]_TaxID=11292;
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192392; AAO45867.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 61
Q80KF2
ID Q80KF2 PRELIMINARY; PRT; 28 AA.
AC Q80KF2;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]_TaxID=11292;
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192390; AAO45865.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 62
Q80KF3
ID Q80KF3 PRELIMINARY; PRT; 28 AA.
AC Q80KF3;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.

```

```

OX NCBI_TaxID=11292;
RN [1]_TaxID=11292;
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192389; AAO45864.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 63
Q80KF4
ID Q80KF4 PRELIMINARY; PRT; 28 AA.
AC Q80KF4;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]_TaxID=11292;
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192388; AAO45863.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 64
Q80KF5
ID Q80KF5 PRELIMINARY; PRT; 28 AA.
AC Q80KF5;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]_TaxID=11292;
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192387; AAO45862.1; -.

```

```
FT NON_TER 28
SQ SEQUENCE 28 AA; 2938 MW; 2B2237A9AB8F3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
    |||||
Db 21 PRGAP 25

RESULT 65
Q80KF6 PRELIMINARY; PRT; 28 AA.
AC Q80KF6;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192386; AAO45861.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
    |||||
Db 21 PRGAP 25

RESULT 66
Q80KF8 PRELIMINARY; PRT; 28 AA.
AC Q80KF8;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192384; AAO45859.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
    |||||
```

```
Db 21 PRGAP 25

RESULT 67
Q80KF9 PRELIMINARY; PRT; 28 AA.
AC Q80KF9;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192383; AAO45858.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
    |||||
Db 21 PRGAP 25

RESULT 68
Q80KG1 PRELIMINARY; PRT; 28 AA.
AC Q80KG1;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192381; AAO45856.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
    |||||
Db 21 PRGAP 25

RESULT 69
Q80KG2 PRELIMINARY; PRT; 28 AA.
AC Q80KG2;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
```

```

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192378; AAO45853.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 70
Q80KG3
ID Q80KG3 PRELIMINARY; PRT; 28 AA.
AC Q80KG3;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192379; AAO45854.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 71
Q80KG4
ID Q80KG4 PRELIMINARY; PRT; 28 AA.
AC Q80KG4;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;

```

```

RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192378; AAO45853.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 72
Q80KG7
ID Q80KG7 PRELIMINARY; PRT; 28 AA.
AC Q80KG7;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192375; AAO45850.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 2982 MW; 2E2927A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAP 5
Db 21 PRGAP 25

RESULT 73
Q80KG8
ID Q80KG8 PRELIMINARY; PRT; 28 AA.
AC Q80KG8;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats.";
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192374; AAO45849.1; -.
FT NON TER 28
SQ SEQUENCE 28 AA; 2998 MW; 2E3147A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;

```


Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Job time : 95.5 secs

Qy 1 PRGAP 5
Db 21 PRGAP 25

RESULT 74

Q80KG9 PRELIMINARY; PRT; 28 AA.
AC Q80KG9;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats."
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192373; AAO45848.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRGAP 5
Db 21 PRGAP 25

RESULT 75

Q80KH0 PRELIMINARY; PRT; 28 AA.
AC Q80KH0;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE L protein (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22541850; PubMed=12655080;
RA Paez A., Nunez C., Garcia C., Boshell J.;
RT "Molecular epidemiology of rabies epizootics in Colombia: evidence for
RT human and dog rabies associated with bats."
RL J. Gen. Virol. 84:795-802(2003).
DR EMBL; AY192372; AAO45847.1; -.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3010 MW; 2E3637A9AACAC3C6A CRC64;

Query Match 25.0%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRGAP 5
Db 21 PRGAP 25

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:46:31 ; Search time 85 Seconds
(without alignments)
84.407 Million cell updates/sec

Title: US-10-066-965A-4

Perfect score: 20

Sequence: 1 PRGAPWLRVCVQMLETKFL 20

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2002273 seqs, 358729299 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	30.0	106	3	AB11672 A. vitis
2	6	30.0	147	2	AAY76621 Human ova
3	6	30.0	169	7	ABO62427 Klebsiell
4	6	30.0	285	5	ABP42001 Human ova
5	6	30.0	315	4	ABB60668 Drosophil
6	6	30.0	395	6	ADA55451 Human pro
7	6	30.0	395	6	AAE33476 Human AGT
8	6	30.0	395	7	ADJ70703 Human hea
9	6	30.0	434	6	ABU49025 Protein e
10	6	30.0	444	4	ABB66870 Drosophil
11	6	30.0	444	4	ABB61409 Drosophil
12	6	30.0	501	4	AAAG5822 Human GPR
13	6	30.0	571	6	ABR58452 Human NOV
14	6	30.0	577	6	ABR58460 Human NOV
15	6	30.0	731	7	ADJ71063 Human hea
16	6	30.0	847	8	ADN99559 Novel hum
17	6	30.0	1290	7	ADL15079 Human adu
18	6	30.0	1490	6	ABR58454 Human NOV
19	6	30.0	1545	6	ABR58453 Human NOV
20	6	30.0	1549	6	ABR58455 Human NOV
21	6	30.0	1551	4	AAAM1081 Human pol
22	6	30.0	1568	8	ADO00979 Mouse hom
23	6	30.0	1588	4	AAB20155 Secreted
24	6	30.0	1588	4	AAAM39295 Human pol
25	6	30.0	1588	6	ABR58450 Human NOV

26	6	30.0	1588	6	ABR58456 Human NOV
27	6	30.0	1588	6	ABR58459 Human NOV
28	6	30.0	1588	6	ABR58462 Human NOV
29	6	30.0	1588	7	ADG38871 Human SEC
30	6	30.0	1588	7	ADG38832 Human SEC
31	6	30.0	1624	5	AAE23379 Human LP2
32	5	25.0	9	2	AAAR37732 Collagen-
33	5	25.0	9	2	AAAR32339 Collagen-
34	5	25.0	9	2	AAW57686 Collagen-
35	5	25.0	9	5	AAU71713 Human MHC
36	5	25.0	9	5	AAU71607 Human MHC
37	5	25.0	10	5	AAU71662 Human MHC
38	5	25.0	12	6	ABP72812 Human can
39	5	25.0	15	2	AAAR93209 New contr
40	5	25.0	15	2	AAW96728 ENA-78 de
41	5	25.0	15	6	ABU67730 Human ang
42	5	25.0	15	7	ABO07364 Angiogene
43	5	25.0	16	2	AAAR91628 Alpha che
44	5	25.0	20	8	ADK01531 Hepatitis
45	5	25.0	22	5	ABG69986 Rabbit pl
46	5	25.0	22	5	ABG69985 Rabbit pl
47	5	25.0	25	8	ADM32132 EphB4 blo
48	5	25.0	26	2	AAAY20230 Human bet
49	5	25.0	28	2	AAAY19651 SEQ ID NO
50	5	25.0	33	3	AAAG56952 Arabidops
51	5	25.0	40	5	ABG69992 Rabbit pl
52	5	25.0	42	5	ADF94734 Hepatitis
53	5	25.0	42	5	ADF94732 Hepatitis
54	5	25.0	44	2	AAW57222 Targeting
55	5	25.0	44	6	AAE38084 Human cyc
56	5	25.0	46	3	AAAB39249 Gene 11 h
57	5	25.0	46	4	AAAM21142 Peptide #
58	5	25.0	46	4	AAAM15409 Peptide #
59	5	25.0	46	4	ABBA43459 Peptide #
60	5	25.0	46	4	ABBA34414 Peptide #
61	5	25.0	46	4	AAAM27897 Peptide #
62	5	25.0	46	4	AAAM37347 Peptide #
63	5	25.0	46	4	ABBA29251 Peptide #
64	5	25.0	46	4	AAU17884 Novel hum
65	5	25.0	46	4	ABBI19826 Protein #
66	5	25.0	46	4	AAAM67599 Human bon
67	5	25.0	46	4	AAAM64389 Human bra
68	5	25.0	46	4	ABG58835 Human liv
69	5	25.0	46	4	ABG49246 Human liv
70	5	25.0	46	4	AAAM03170 Peptide #
71	5	25.0	46	5	ABG46222 Human pep
72	5	25.0	46	5	ABG37192 Human pep
73	5	25.0	46	7	ADGA1264 Human res
74	5	25.0	47	4	ABBA38152 Peptide #
75	5	25.0	47	4	AAAM31580 Peptide #
76	5	25.0	47	4	ABBA23358 Protein #
77	5	25.0	47	4	AAAM71301 Human bon
78	5	25.0	47	4	AAAM58788 Human bra
79	5	25.0	47	4	ABG53012 Human liv
80	5	25.0	47	5	ABGA1101 Human pep
81	5	25.0	50	4	ABGI19401 Novel hum
82	5	25.0	50	7	ADG85160 Tobacco c
83	5	25.0	50	8	ADN10509 Nicotiana
84	5	25.0	50	8	ADG59076 Tobacco c
85	5	25.0	54	3	AAAG57910 Arabidops
86	5	25.0	54	3	AAAG61712 Arabidops
87	5	25.0	54	6	ABU62915 Mouse ost
88	5	25.0	55	3	AAAG60769 Arabidops
89	5	25.0	55	5	ABP08831 Human ORF
90	5	25.0	57	5	ABP64607 Human ORF
91	5	25.0	59	4	AAAO0361 Human pol
92	5	25.0	60	4	AAAM88090 Human imm
93	5	25.0	62	4	AAU46690 Propionib
94	5	25.0	62	6	ABMA43209 Propionib
95	5	25.0	63	3	AAAG50768 Arabidops
96	5	25.0	63	8	ABO58509 Human gen
97	5	25.0	64	5	ABP00767 Human ORF
98	5	25.0	65	3	AAAG61788 Arabidops

99 5 25.0 66 5 ADF94731 Adf94731 Hepatitis
100 5 25.0 66 5 ADF94733 Adf94733 Hepatitis

ALIGNMENTS

RESULT 1
AAB11672
ID AAB11672 standard; protein; 106 AA.
XX
AC AAB11672;
XX
DT 23-OCT-2000 (first entry)
XX
DE A. vitis hypersensitive response elicitor protein, SEQ ID NO:59.
XX
KW Hypersensitive response elicitor protein; HR; disease resistance;
KW insecticide; fungicide; antiviral; bactericide; growth enhancer;
KW stress resistance; transgenic plant.
XX
OS Agrobacterium vitis.
XX
PN WO200028056-A2.
XX
PD 18-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US026079.
XX
PR 06-NOV-1998; 98US-0107387P.
XX
PA (CORR) CORNELL RES FOUND INC.
XX
PI Burr TJ, Herlache TC, Zhang H;
XX
DR WPI; 2000-376567/32.
DR N-PSDB; AAA61517.
XX
PT New protein from Agrobacterium vitis, useful e.g. for imparting
PT resistance to disease or stress to plants, is involved in production of a
PT hypersensitive response.
XX
PS Claim 3; Page 131; 157pp; English.
XX
CC Sequences AAB11630-B11688 represent proteins from Agrobacterium vitis
CC which elicit a hypersensitivity response (HR) in a plant. The invention
CC also relates to nucleotide sequences (AAA61501-A61524) encoding the A.
CC vitis HR elicitor proteins. The HR is a rapid, localised necrosis that is
CC associated with the active defence of plants against many pathogens, and
CC occurs when a pathogenic organism interacts with a nonhost plant (i.e.
CC one in which intracellular bacterial growth and disease development do
CC not occur). Like other HR elicitors, the A. vitis elicitor functions in
CC non-host plants by causing a rapid hypersensitive response that results
CC in walling-off and killing of the pathogen. On grape plants, the A. vitis
CC elicitor induces a restricted necrosis of tissues, resulting in the death
CC of plant cells and induction of pathogen resistance. A. vitis HR elicitor
CC proteins, in non-infectious form, are used to treat plants or their seeds
CC to impart resistance to disease, such as those caused by fungi, bacteria
CC or viruses; and to enhance growth, e.g., to increase yield or to provide
CC earlier germination or maturation. The proteins can also be used to
CC control insects, to impart resistance to environmental stresses, e.g.,
CC cold, and to improve nutritional value, e.g., altered oil content. The
CC same effects can be produced by producing transgenic plants or seeds by
CC incorporation of DNA that encodes A. vitis HR elicitor proteins. Use of
CC A. vitis HR elicitor proteins, or nucleic acids encoding them, may allow
CC control of previously untreatable diseases; provide systemic treatment;
CC and eliminate the need for biological control agents or polluting
CC chemicals
XX
SQ Sequence 106 AA;
Query Match 30.0%; Score 6; DB 3; Length 106;
Best Local Similarity 100.0%; Pred. No. 80;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 APMWLR 9
DB 82 APMWLR 87
RESULT 2
AAY76621
ID AAY76621 standard; protein; 147 AA.
XX
AC AAY76621;
XX
DT 10-APR-2000 (first entry)
XX
DE Human ovarian tumor EST fragment encoded protein 117.
XX
KW Expressed sequence tag; EST; human; ovarian tumor; anticancer;
KW gene therapy; treatment.
XX
OS Homo sapiens.
XX
PN DE19817557-A1.
XX
PD 21-OCT-1999.
XX
PF 09-APR-1998; 98DE-01017557.
XX
PR 09-APR-1998; 98DE-01017557.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX
DR WPI; 1999-591920/51.
DR N-PSDB; AAZ77497.
XX
PT New nucleic acid sequences expressed in ovarian, and some other, cancer
PT tissues, and derived polypeptides, for treatment of ovarian cancer and
PT identification of therapeutic agents.
XX
PS Claim 25; Page 292; 310pp; German.
XX
CC This invention describes novel nucleic acid (cDNA) sequences (A) which
CC have anticancer activity and are highly expressed in ovarian tumor tissue
CC (and some also in testis and breast cancer tissue). The products of the
CC invention can be used for gene therapy. (A) are used (i) for recombinant
CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)
CC are used (i) to identify agents suitable for treatment of ovarian cancer;
CC (ii) directly for treating this form of cancer (including expression from
CC gene therapy vectors) and (iii) for generation of specific antibodies.
CC (A) are identified by assembling ESTs (expressed sequence tags) from a
CC particular tissue type before comparison of expression patterns. This
CC allows a significantly longer fragment of the gene to be revealed, so
CC should reduce the number of failures associated with the fact that ESTs
CC from different libraries may represent different parts of the same
CC unknown gene, distorting the estimated frequency of occurrence in a
CC particular tissue. AAY76505-Y76638 represent protein fragments encoded by
CC the human ovarian tumor cDNA library derived EST fragments represented in
CC AAZ77450-Z77572
XX
SQ Sequence 147 AA;
Query Match 30.0%; Score 6; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PRGAPM 6
DB 27 PRGAPM 32
RESULT 3

```
ABO62427
ID ABO62427 standard; protein; 169 AA.
XX AC ABO62427;
XX DT 29-JUL-2004 (first entry)
XX DE Klebsiella pneumoniae polypeptide seqid 8944.
XX KW Recombinant expression vector; transcription regulatory element;
XX KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX OS Klebsiella pneumoniae.
XX PN US6610836-B1.
XX PD 26-AUG-2003.
XX PF 27-JAN-2000; 2000US-00489039.
XX PR 29-JAN-1999; 99US-0117747P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton GL, Osborne M;
XX WPI: 2003-895346/82.
XX N-PSDB; ACH95978.
XX PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX PT preparing a vaccine composition against Klebsiella pneumoniae.
XX PS Disclosure; SEQ ID NO 8944; 932pp; English.
XX CC The invention describes a new isolated nucleic acid encoding a Klebsiella
XX CC pneumoniae polypeptide. Also described are: a recombinant expression
XX CC vector comprising the nucleic acid, operably linked to a transcription
XX CC regulatory element; and a cell comprising the recombinant expression
XX CC vector. The nucleic acid is useful for preparing a vaccine composition
XX CC against Klebsiella pneumoniae. This is the amino acid sequence of a
XX CC Klebsiella pneumoniae polypeptide of the invention
XX SQ Sequence 169 AA;
Query Match 30.0%; Score 6; DB 7; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
QY 3 GAPMWL 8
Db 90 GAPMWL 95
RESULT 4
ABP42001
ID ABP42001 standard; protein; 285 AA.
XX AC ABP42001;
XX DT 22-AUG-2002 (first entry)
XX DE Human ovarian antigen HTELJ95, SEQ ID NO:3133.
XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
XX KW inflammatory condition; immune disorder; blood disorder;
XX KW cardiovascular disorder; respiratory disorder; neurological disorder;
XX KW gastrointestinal disorder; urinary system disorder; drug screening;
XX KW gene therapy; chromosome mapping; forensic analysis;
XX KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX KW antiinflammatory; gynaecological; reproductive.
```

```
XX OS Homo sapiens.
XX PN WO200200677-A1.
XX PD 03-JAN-2002.
XX PF 07-JUN-2001; 2001WO-US018569.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Birse CE, Rosen CA;
XX WPI: 2002-147878/19.
XX N-PSDB; ABQ55078.
XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
XX PT cancer), immune disorders, cardiovascular disorders and neurological
XX PT diseases.
XX PS Claim 11; SEQ ID NO 3133; 2922pp; English.
XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX CC encompasses polypeptides 90% identical and polynucleotides 95% identical
XX CC to the sequences of the invention. The invention additionally relates to
XX CC recombinant vectors and host cells comprising human ovarian antigen
XX CC polynucleotides, antibodies against human ovarian antigens, and the use
XX CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX CC treating, prognosing or preventing various ovary and/or breast-related
XX CC disorders. Such conditions include ovarian cancer and breast cancer, and
XX CC metastatic tumours of ovarian or breast origin, reproductive system
XX CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
XX CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX CC vaginitis), immune disorders (e.g., congenital and acquired
XX CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX CC respiratory disorders, neurological disorders, gastrointestinal disorders
XX CC and urinary system disorders. Ovarian antigen polypeptides and
XX CC polynucleotides may also be used in screening for compounds which
XX CC modulate ovarian antigen expression or activity. The polynucleotides may
XX CC further be used for gene therapy, chromosome mapping, in the
XX CC identification of individuals and in forensic analysis, and the
XX CC polypeptides may be used as food additives or to prepare antibodies
XX CC useful in disease diagnosis, drug targeting and phenotyping. The present
XX CC sequence represents a human ovarian antigen of the invention. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 285 AA;
Query Match 30.0%; Score 6; DB 5; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
QY 15 LETKFL 20
Db 178 LETKFL 183
RESULT 5
ABB60668
ID ABB60668 standard; protein; 315 AA.
XX AC ABB60668;
XX DT 26-MAR-2002 (first entry)
```

XX DE Drosophila melanogaster polypeptide SEQ ID NO 8796.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX DR N-PSDB; ABL04771.
XX CC New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX PS Disclosure; SEQ ID NO 8796; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABE57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 315 AA;

Query Match 30.0%; Score 6; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LETKFL 20
Db 40 LETKFL 45
|||||

RESULT 6
ADA55451
ID ADA55451 standard; protein; 395 AA.
XX AC ADA55451;
XX DT 20-NOV-2003 (first entry)
XX DE Human protein, SEQ ID 3019.
XX KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.
XX OS Homo sapiens.
XX PN EP1293569-A2.
XX PD 19-MAR-2003.
XX PF 21-MAR-2002; 2002EP-00006586.

XX PR 14-SEP-2001; 2001JP-00328381.
XX PR 24-JAN-2002; 2002US-0350435P.
XX PA (HELI-) HELIX RES INST.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Naganari K, Masuho Y;
XX DR WPI; 2003-395539/38.
XX DR N-PSDB; ADA53812.
XX PT New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX PS Claim 14; SEQ ID NO 3019; 205pp; English.
XX CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX SQ Sequence 395 AA;

Query Match 30.0%; Score 6; DB 6; Length 395;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 QMLETK 18
Db 365 QMLETK 370
|||||

RESULT 7
AAE34476
ID AAE34476 standard; protein; 395 AA.
XX AC AAE34476;
XX DT 14-MAY-2003 (first entry)
XX DE Human AGT-114 protein.
XX KW Obesity; anorexia; weight maintenance; impaired muscle development;
KW diabetes; alkylguanine alkyltransferase; energy imbalance; enzyme;
KW gene therapy; human; AGT.
XX OS Homo sapiens.
XX PN WO200295020-A1.
XX PD 28-NOV-2002.
XX PF 21-MAY-2002; 2002WO-AU000628.
XX PR 21-MAY-2001; 2001AU-00005137.
XX PA (AUTO-) AUTOGEN RES PTY LTD.
PA (UYDE-) UNIV DEAKIN.
PA (ITDI-) INT DIABETES INST.
XX PI Collier G, Walder K, Miller JE;
XX DR WPI; 2003-140372/13.
XX PT New isolated nucleic acid molecule expressed in liver or stomach tissue,
PT useful for diagnosing or treating obesity, anorexia, diabetes or energy
PT imbalance, and as targets for agents which act as modulators of
PT physiological processes.